UNIVERSIDADE PAULISTA - UNIP

PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM ENGENHARIA DE PRODUÇÃO

LILIAM SAYURI SAKAMOTO

OTIMIZAÇÃO DA DETECÇÃO DE PERDA DE DADOS ATRAVÉS DE *DATA LOSS PREVENTION* – DLP E ALGORITMOS BASEADOS NA LÓGICA PARACONSISTENTE ANOTADA EVIDENCIAL Eτ

SÃO PAULO

2023

UNIVERSIDADE PAULISTA - UNIP

PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM ENGENHARIA DE PRODUÇÃO

OTIMIZAÇÃO DA DETECÇÃO DE PERDA DE DADOS ATRAVÉS DE *DATA LOSS PREVENTION* – DLP E ALGORITMOS BASEADOS NA LÓGICA PARACONSISTENTE ANOTADA EVIDENCIAL Eτ

Tese de Doutoramento apresentada ao Programa de Doutorado em Engenharia de Produção da Universidade Paulista - UNIP como requisito parcial para obtenção do grau de Doutor em Engenharia

Orientador: Prof. Dr. Jair Minoro Abe

Área de Concentração: Gestão de Sistemas de Operação

Linha de Pesquisa: Métodos Quantitativos em Engenharia de Produção

Projeto de Pesquisa: Métodos Quantitativos, Computacionais e Tecnológicos em Engenharia de Produção

LILIAM SAYURI SAKAMOTO

SÃO PAULO

2023

Sakamoto, Liliam Sayuri.
Otimização da detecção de perda de dados através de *Data Loss Prevention* – DLP e algoritmos baseados na lógica paraconsistente anotada evidencial Eτ / Liliam Sayuri Sakamoto. – 2023.
184 f. : il. color. + CD-ROM.

Tese de Doutorado Apresentada ao Programa de Pós Graduação em Engenharia de Produção da Universidade Paulista, São Paulo, 2023.
Área de concentração: Gestão de Sistemas de Operação.
Orientador: Prof. Dr. Jair Minoro Abe.

1. DLP – *Data Loss Prevention*. 2. Lógica paraconsistente anotada evidencial Eτ. 3. DLP paraconsistente. I. Abe, Jair Minoro (orientador). II. Título.

Ficha elaborada pelo Bibliotecário Rodney Eloy CRB8-6450

Dedicatória

Dedico ao meu marido Moacir, pelo seu amor e por fazer todas as coisas para que esta tese se tornasse possível.

AGRADECIMENTOS

Agradeço:

À Deus por conseguir chegar até esta etapa desse estudo.

Meu Orientador Prof. Dr. Jair Minoro Abe por ser decisivo neste estudo e por compreensão e auxílio neste momento de minha vida.

Em especial o amigo que sempre me apoiou em todos os passos dessa jornada o Prof. Dr. Luiz Antônio de Lima.

Aos Professores que marcaram este curso com suas aulas que só contribuíram ainda mais para meu aprofundamento pessoal e profissional, Profa. Irenilza, Prof. Oduvaldo e Prof. Pedro.

Aos colegas Nilson Amado de Souza, Jonatas de Souza, Angel Martinez, Aparecido Carlos Duarte, entre outros.

LILIAM SAYURI SAKAMOTO

OTIMIZAÇÃO DA DETECÇÃO DE PERDA DE DADOS ATRAVÉS DE DATA LOSS PREVENTION – DLP E ALGORITMOS BASEADOS NA LÓGICA PARACONSISTENTE ANOTADA EVIDENCIAL Eτ

Tese apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Paulista – UNIP, para a obtenção do título de Doutor em Engenharia de Produção.

Orientador: Prof. Dr. Jair Minoro Abe

Aprovado em:

BANCA EXAMINADORA:

_____/__/___

Prof. Dr. Jair Minoro Abe

Universidade Paulista – UNIP

_____/__/___

Prof. Dra. Irenilza de Alencar Nääs

Universidade Paulista – UNIP

_____/__/___

Prof. Dr. Marcelo Okano

Universidade Paulista UNIP

_____/__/___

Prof. Dr. João Inácio da Silva Filho - UNISANTA

_____/__/___

Prof. Dr. Cláudio Rodrigo Torres - FATEC - SP

Suplentes: Prof. Dr. Oduvaldo Vendrametto – UNIP

Prof. Dr. Alessandro Campolina – ICESP/HC – USP

RESUMO

O foco deste estudo é otimizar a detecção do nível de perda de dados. Muitas empresas ao redor do mundo possuem uma lacuna no controle a organização de dados estruturados e não estruturados. Todos esses dados podem ser agrupados em repositórios, porém, a análise sobre perda de dados, ou seja, prevenção de vazamento de dados, com a quebra da privacidade, direciona o uso de alguns critérios de Inteligência Artificial com uso de *DLP – Data Loss Prevention*. Este trabalho se fundamenta na aplicação da Lógica Paraconsistente Anotada Evidencial Eτ, através do uso em conjunto com o DLP. A técnica de análises baseada em um DLP de mercado utiliza modelos padrões que são monitorados de forma limitada, e frequentemente apresentam uma massa de contradições e falhas, levando a tomadas de decisões com falsos positivos. Inicialmente foi utilizado a programação *Python*, para formar uma estrutura do DLP funcionando em conjunto com algoritmos paraconsistentes, o que foi denominado DLP Paraconsistente. O DLP Paraconsistente foi aplicado em testes com dados de uma empresa transportadora, analisando seus incidentes que apresenta 60% de deteção de perda de dados. Na pesquisa desenvolvida nesta tese é feita uma comparação entre DLP convencional (dados fornecidos pela empresa) e a DLP Paraconsistente com o objetivo de otimizar as análises para a minimização da perda de dados.

Palavras-chave: *DLP – Data Loss Prevention*, Lógica Paraconsistente Anotada Evidencial Eτ, DLP Paraconsistente.

ABSTRACT

The focus of this study is to optimize this analysis while minimizing the level of data loss. Many companies around the world have a gap in controlling the organization of structured and unstructured data. All this data can be grouped into repositories, however, the analysis of data loss, that is, prevention of data leakage, with the breach of privacy, directs the use of some Artificial Intelligence criteria with the use of DLP – Data Loss Prevention. This work is based on the application of the Annotated Eτ Paraconsistent Logic, through use in conjunction with DLP. The analysis technique based on a market DLP uses standard models that are monitored in a limited way, and often present a mass of contradictions and flaws, leading to false positive decision-making. Initially, Python programming was used to form a DLP structure working in conjunction with paraconsistent algorithms, which was called Paraconsistent DLP. The Paraconsistent DLP was applied in tests with data from a transport company, analyzing its incidents, which showed a 60% data loss detection rate. In the research developed in this thesis, a comparison is made between conventional DLP (data provided by the company) and Paraconsistent DLP with the aim of optimizing analyses to minimize data loss.

Keywords: DLP – Data Loss Prevention, Paraconsistent Logic Evidential Annotated Eτ, Paraconsistent DLP.

UTILIDADE

Segundo o artigo nº46 da Lei nº 13.709/2018 – Lei Geral de Proteção de Dados Pessoais os agentes de tratamento (Controladores e Operadores) devem adotar medidas de segurança técnicas e administrativas de acesso não autorizado e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou qualquer forma de tratamento inadequado ou ilícito, que direciona o estudo desenvolvido.

Dos 17 Objetivos de Desenvolvimento Sustentável (ODS) das Nações Unidas, um deles está diretamente ligado ao estudo, que é o 9º objetivo: Indústria, Inovação e infraestrutura, com intuito de construir infraestrutura resiliente, promover a industrialização inclusiva e sustentável, e fomentar a inovação.

Ao se cruzar estas duas informações, se pode elencar alguns fatores granulares, como:

- Identificação e solução de problemas com privacidade de dados;
- Mitigar ou prevenir riscos;
- Assegurar reação rápida e eficaz para ocorrência de catástrofe;
- Reforço das capacidades de resiliência com melhoria contínua.

Dada a importância e a atualidade das questões estudadas e apresentadas no presente trabalho, entende-se como válida a contribuição, como ferramenta para o auxílio dos agentes definidos na LGPD denominados “encarregado de dados”.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

LISTA DE ABREVIATURAS

ANPD – Agência Nacional de Proteção de Dados

APDADOS – Associação Nacional do Profissionais de Privacidade de Dados

CVE - Listas de Segurança Pública ou *Common Vulnerabilities and Exposures*

DC – Dados Corporativos

DLP – *Data Loss Prevention*

DP – Dados Pessoais

DPO – *Data Protection Officer*

GDPR – Regulamento Geral de Proteção de Dados na União Europeia

IA – Inteligência Artificial

LGPD – Lei Geral de Privacidade de Dados

MITRE ATT e CK - Diretriz para classificar e descrever ataques cibernéticos e invasões

NIST - Instituto Nacional de Padrões e Tecnologia

SIEM - *Security Information and Event Monitoring*

SOC – *Security Operation Center*

TOP 10 OWASP - Guia para testes de segurança em aplicações *web*

LISTA DE SÍMBOLOS

V	Verdadeiro
F	Falso
T	Inconsistente
$\perp$	Paracompleto
QV $\rightarrow$ T	Quase Verdadeiro tendendo ao Inconsistente
QV $\rightarrow \perp$	Quase Verdadeiro tendendo ao Paracompleto
QF $\rightarrow$ T	Quase Falso tendendo ao Inconsistente
QF $\rightarrow \perp$	Quase Falso tendendo ao Paracompleto
QI $\rightarrow$ V	Quase Inconsistente tendendo ao Verdadeiro
QI $\rightarrow$ F	Quase Inconsistente tendendo ao Falso
Q$\perp \rightarrow$ V	Quase Paracompleto tendendo ao Verdadeiro
Q$\perp \rightarrow$ F	Quase Paracompleto tendendo ao Falso
x_1	Dados estratégicos
x_2	Dados táticos
x_3/x_n	Dados operacionais
Wk_1	Peso 3
Wk_2	Peso 2
Wk_3/Wk_p	Peso 1
b_k	Bias
Σ	Função somatória
ϕ (.)	Função de ativação para Lógica Et
u_k	Valor para liberar a função de ativação
y_k	Saída (percentuais de dados considerados válidos ou inconsistentes)

Sumário

CAPÍTULO 1 – INTRODUÇÃO

1.1 Contexto Atual

Atualmente, em todos os países há a preocupação com a privacidade de dados pessoais, em vários deles inúmeras legislações a respeito surgiram, como o Regulamento Geral de Proteção de Dados na União Europeia - GDPR (*European Congress*, 2022). No Brasil ocorreu com a entrada em vigor da Lei Geral de Privacidade de Dados nº 13.709/2018 – LGPD (BRASIL, 2022). Atualmente se enfatiza o quão é importante não somente para a pessoa física e sua integridade a proteção e a privacidade de seus dados, que hoje são negociados até no mercado negro, pois um CPF ou dados de cartão de crédito podem ser usados para empréstimos e financiamentos, enquanto para uma corporação uma credencial ou matrícula, por exemplo, de um funcionário pode dar acesso a muitos sistemas de impacto financeiro e até de mercado.

Dentro de uma empresa a quantidade de informações pessoais que são capturadas, guardadas, transformadas em relatórios e eliminadas em um dia já chega a números exponenciais, que só aumentam com o passar do tempo, isso para os dados estruturados e identificados, ou seja, os que estão em Sistemas Aplicativos, Sistemas Especialistas e Banco de Dados (HELLEBRAND,2017).

Entretanto, existem outros tipos de dados, os não estruturados que transitam nos aplicativos de Redes Sociais, dentro de conteúdos de e-mails que geralmente não analisados pelas empresas, ou armazenados em *pendrive* ou *HD* externo. Este tipo dado é mais complexo de se identificar, e muitas vezes esse conteúdo não é devidamente analisado, o que pode ocasionar a perda dessas informações, um vazamento por falta de guarda adequada, ou até extração indevida por pessoas não autorizadas por falta de classificação do nível de acesso ao dado (EBERENDU, 2016).

Com a necessidade de ferramentas para assertividade e aderência a LGPD, que auxiliem na função do *Data Protection Officer* - DPO, que pela Lei é denominado no Brasil como encarregado de dados. Entretanto, as normativas são incisivas quanto ao uma nova abordagem na captação de dados pessoais pelas empresas de seus clientes, funcionários, fornecedores e prestadores de

serviço, devendo seguir o Regimento de Dosimetria publicado pela Agência Nacional de Proteção de Dados - ANPD - fevereiro de 2023, que descreve os níveis de multas a serem aplicadas, desde advertências simples até situações em que pode existir a paralisação total das atividades da empresa, conforme o tipo e quantidade de incidentes ou reclamações ocorridas (ANPD, 2023).

1.2 Objetivo

1.2.1 Objetivo Geral

O objetivo geral desta pesquisa é estudar a otimização da detecção de perda de dados através de uma estrutura algorítmica denominada *Data Loss Prevention – DLP* Paraconsistente que utiliza técnica prevenção de perda de dados e algoritmos baseados na Lógica Paraconsistente Anotada Evidencial $E\tau$ - Lógica $E\tau$ para dano/roubo/vazamento de dados pessoais não estruturados com foco na Lei Geral de Proteção de Dados – LGPD, para auxílio aos encarregados de dados ou *Data Protection Officer - DPO*.

1.2.2 Objetivos Específicos

- Prover identificação de perda/dano/roubo/vazamento de dados ou *Data Loss Prevention – DLP* Paraconsistente:
    - Identificar perda de dados pessoais não estruturados (clientes, funcionários e fornecedores - incidentes LGPD);
    - Identificar perda de dados corporativos não estruturados (rede, armazenamento, *endpoint* e *cloud*);
- Detalhar a estrutura do *Data Loss Prevention* - DLP Paraconsistente;
- Aplicar *Data Loss Prevention* - DLP Paraconsistente para análise de perda/dano/roubo/vazamento de dados não estruturados.

1.3 Justificativa

Conforme o Guia de Privacidade da Organização para a Cooperação e Desenvolvimento Econômico – OCDE de 2013, os dados pessoais já passavam a desempenhar um papel cada vez mais importante nas economias e as sociedades, além da importância na vida de cada indivíduo. Visto que as inovações nas tecnologias da informação e na comunicação proporcionaram

impacto nas operações empresariais, na administração governamentais, bem como nas atividades individuais. Estas novas tecnologias em conjunto com a utilização responsável dos dados produzem grandes benefícios sociais e econômicos (OCDE, 2013).

É notório que o volume de dados pessoais recolhidos, utilizados e armazenados é muito vasto e continua a crescer diariamente. As redes de comunicações modernas, como a Internet apoiam a acessibilidade global, potencializando o aumento dos dados pessoais com resultado de suas análises e outras informações abrangentes sobre o direcionamento, interesses e atividade dos indivíduos, ou seja, os dados são cada vez mais utilizados de formas não previstas do momento da sua captação. Esse é um dos motivos que levaram a necessidade de regulamentações para que não haja riscos para a privacidades dos indivíduos (HORODYSKI, 2014) e (OCDE, 2013).

Os Gestores de áreas de Engenharia de Produção entre outras, precisam estar preparados as novas exigências tecnológicas, corporativas e regulatórias. Hoje em dia, não é mais uma função unicamente da área de Tecnologia de Informação dominar técnicas e métodos inovativos para otimização de processos operacionais. Em empresas que trabalham com enorme quantidade de dados (estruturados e não estruturados), se torna um desafio constante a manutenção dessas bases e consolidação dessas informações diariamente. Pode-se visualizar em termos crescentes a busca por ferramentais que alinhem a experiência de profissionais especializados com aplicações com capacidade de cognição próprios do homem (Inteligência Artificial), porém implantados em aplicações como *Data Loss Prevention – DLP* Paraconsistente que poderão automatizar e otimizar tomadas de decisões em situações duvidosas, incertas e complexas com foco em privacidade de dados pessoais com base na LGPD.

1.4 Motivação do trabalho

Aplicação do desenvolvimento de Inteligência Artificial, sendo que ferramentas automatizadas do mercado para detecção de incidentes de privacidade de dados baseiam-se na maioria dos casos na lógica clássica e apresentam muitos falsos-positivos, sua calibração é demorada, sendo que os

resultados obtidos não são satisfatórios de imediato, mas após várias calibrações. São elaboradas principalmente para análise de *logs* de rede, de *firewall*, movimentação de usuário e não especificamente para privacidade de dados não estruturados, como os que constam de aplicativos de rede sociais, e-mails, físicos (impressos em papel) ou apresentados em *display* ou somente em tela de Sistemas Aplicativos.

A ferramenta proposta visa uma otimização para que as informações capturadas sobre supostos incidentes envolvendo dados não estruturados, sejam analisadas por um *Data Loss Prevention – DLP* Paraconsistente, e que se possa trazer evidenciação refinada sobre quais são os reais incidentes de privacidade de dados e quais são falsos positivos, quanto a perda/dano/vazamento desses dados pessoais.

Visto que, a maior parte das ferramentas *Security Information and Event Monitoring* - SIEM de mercado traz diariamente cerca de 1.000 a 2.000 dados para serem analisados, geralmente referente a questões de cibersegurança, mas um encarregado de dados irá precisar analisar individualmente quais dizem respeito as questões reais de incidentes de privacidade de dados, esta demanda justifica o estudo. Observando-se que em grandes organizações esta média chega a ser exponencial (CINQUE, 2018).

1.5 Estrutura do Trabalho

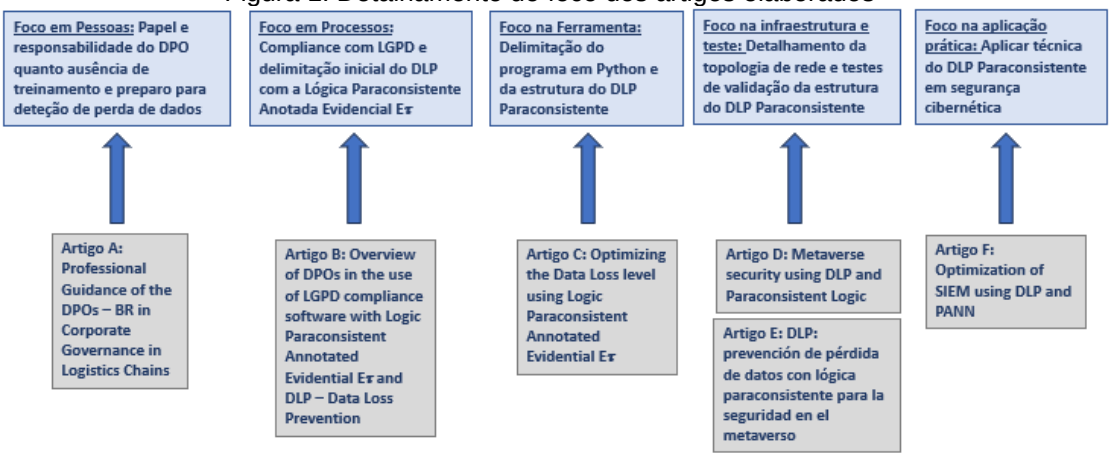
Os capítulos foram divididos em:

- No Capítulo 1 apresenta-se a Introdução, com a contextualização geral, o Objetivo Geral do trabalho, os Objetivos Específicos detalhados, bem como a justificativa baseada nos estudos da Organização para a Cooperação e Desenvolvimento Econômico – OCDE de 2013 e motivação para o desenvolvimento dele;
- No Capítulo 2 é detalhada Metodologia, com a utilização do Sistema de Classificação de Pesquisas, conforme GIL (2018), por:
    - Finalidade como Pesquisa Aplicada de estudo com o *Data Loss Prevention - DLP* Paraconsistente;

- Objetivos gerais como Pesquisa Exploratória;
- Abordagem metodológica como Pesquisa Quantitativa;
- Método empregado:
  - Pesquisa bibliográfica e
  - Pesquisa experimental.

- No Capítulo 3 é feita a Revisão Bibliográfica que aborda: Lógica Paraconsistente Anotada Evidencial Eτ, Proteção e Privacidade de Dados, *Data Loss Prevention – DLP* e Ameaças Cibernéticas;
- No Capítulo 4 é apresentada a elaboração da ferramenta *Data Loss Prevention - DLP* Paraconsistente, com: explicação do Sistema de Gerenciamento de Privacidade de Dados Pessoais, definições e proposições sobre o aspecto regulatório, escolha dos especialistas, fixação de valore de controle, escolha de fator de influência, construção da estrutura do DLP Paraconsistente, definição de camadas internas, construção de Base de Dados, definição da tomada de decisões;
- No Capítulo 5 são apresentados os Resultados Obtidos por meio dos artigos apresentados conforme figura 1:

Figura 1: Detalhamento do foco dos artigos elaborados



Fonte: Autora

- Demonstrar através de *Data Loss Prevention* – DLP Paraconsistente para prevenção de dano/roubo/vazamento

de dados pessoais não estruturados, com foco em pessoas, ou seja, na responsabilidade do encarregado de dados quanto ausência de treinamento e preparo para deteção de perda de dados:

- Artigo A: *Professional Guidance of the DPOs – BR in Corporate Governance in Logistics Chains*, neste artigo aborda a importância de um guia de treinamento para os Encarregados de Dados ou DPOs (ANEXO I), apresentado no Congresso APMS 2022 e publicado no link: https://link.springer.com/chapter/10.1007/978-3-031-16411-8_8. O foco desse artigo é pessoas, ou seja, a importância do desenvolvimento e responsabilização do encarregado de Dados para sua atuação nas empresas de logística como em outras.

- Identificar perda de dados pessoais não estruturados (clientes, funcionários e fornecedores - incidentes LGPD), com foco em processos de compliance com LGPD e delimitação inicial do DLP com a Lógica Paraconsistente Anotada Evidencial Eτ:

  - Artigo B: *Overview of DPOs in the use of LGPD compliance software with Logic Paraconsistent Annotated Evidential Eτ and DLP – Data Loss Prevention*, apresentado e publicado no 9º Congresso da Turquia *ÇUKUROVA* 2022 (ANEXO II), publicado o *book* de *full text* em: https://en.iksadkongre.net/kongre-kitaplari. O foco desse artigo é em processos, principalmente no de Compliance com a LGPD e o desenvolvimento de *software* aderente a Lei.

- Identificar perda de dados corporativos não estruturados (rede, armazenamento, *endpoint* e *cloud*), com foco na

ferramenta com a delimitação do programa em Python e a estrutura do DLP Paraconsistente:

- Artigo C: *Optimizing the Data Loss level using Logic Paraconsistent Annotated Evidential Eτ* (ANEXO III), publicado como Capítulo do Livro *Advances in Applied Logics* pela *Springer* em: https://link.springer.com/chapter/10.1007/978-3-031-35759-6_9. O foco desse artigo foi a ferramenta, o desenvolvimento do Programa em *Python* com a Identificação dos dados válidos e dos inconsistentes e a utilização do Para-analisador para revalidar esses dados e reaproveitá-los, minimizando os riscos da detecção de perda de dados.

- Detalhar a estrutura do *Data Loss Prevention - DLP* Paraconsistente, foco na infraestrutura e teste, com o detalhamento da topologia de rede e validação da estrutura do DLP Paraconsistente:

  - Artigo D: *Metaverse security using DLP and Paraconsistent Logic* (ANEXO IV), submetido para revista A1 - Journal of Management in Engineering. Neste artigo é apresentada a topologia da rede usada no processo, bem como é realizado um teste com uma pequena massa de dados para validar a utilização do DLP Paraconsistente.

  - Artigo E: *DLP: prevención de pérdida de datos con lógica paraconsistente para la seguridad en el metaverso* (ANEXO VI), submetido para revista A1 - *Enseñanza de las Ciencias.* Neste artigo, continuamos com a mesma topología apresentada no artigo anterior, entretanto a amostra testada é maior para validar a utilização da ferramenta DLP Paraconsistente.

- Aplicar *Data Loss Prevention - DLP* Paraconsistente para análise de perda/dano/roubo/vazamento de dados não estruturados, com foco na aplicação prática, na aplicação da técnica do DLP Paraconsistente em segurança cibernética.
    - Artigo F: *Optimization of SIEM using DLP and PANN*, (ANEXO V), submetido para revista A1 – IEEE System Journal. O foco desse artigo é a aplicação prática do DLP Paraconsistente, que apresentou otimização da utilização de um Sistema de Monitoramento de Eventos de Incidentes Cibernéticos.

- No Capítulo 6 são apresentadas as Considerações Finais com detalhamento dos Resultados apresentados nos artigos citados no item anterior, da Discussão sobre dados estruturados e Trabalhos futuros objetivando bases Legais exigidas pela LGPD.

CAPÍTULO 2 – METODOLOGIA

2.1 Procedimento Metodológico

Pode-se, conforme GIL (2018) com base no Sistema de Classificação de Pesquisas classificar este estudo quanto:

- A finalidade como uma Pesquisa Aplicada, pois se pode utilizar de forma imediata os resultados para a situação de detecção de perda de dados;
- Aos objetivos gerais como uma Pesquisa Exploratória, pois visa a proporcionar maior familiaridade com a utilização do DLP Paraconsistente, para torná-lo mais explícito e formalizar a hipótese da otimização da detecção de perda de dados;
- A abordagem metodológica como uma Pesquisa Quantitativa, pois caracteriza-se pelo uso de dados anonimizados para os testes do DLP Paraconsistente e a revalidação dos dados considerados inconsistentes;
- Aos métodos empregados como uma:
    - Pesquisa Bibliográfica com uma revisão da Lógica Eτ, Algoritmo Para-analisador, *Data Loss Prevention - DLP* e Proteção de Dados;
    - Pesquisa Experimental referente a *Data Loss Prevention – DLP* Paraconsistente por meio dos testes realizados nos artigos apresentados nos ANEXOS I a VI:
        - ANEXO I - Artigo A: *Professional Guidance of the DPOs – BR in Corporate Governance in Logistics Chains*;
        - ANEXO II - Artigo B: *Overview of DPOs in the use of LGPD compliance software with Logic Paraconsistent Annotated Evidential Eτ and DLP – Data Loss Prevention*;

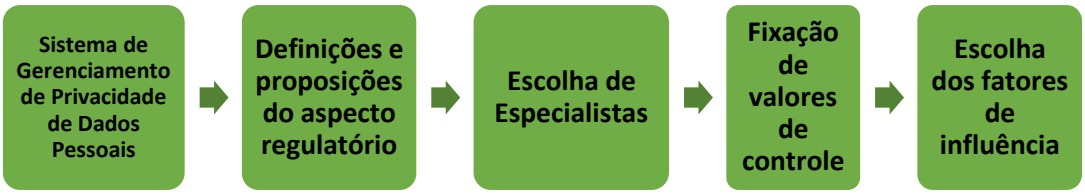
- ANEXO III - Artigo C: *Optimizing the Data Loss level using Logic Paraconsistent Annotated Evidential Eτ*;
- ANEXO IV - Artigo D: *Metaverse security using DLP and Paraconsistent Logic*;
- ANEXO V - Artigo F: *Optimization of SIEM using DLP and PANN*;
- ANEXO VI - Artigo E: *DLP: prevención de pérdida de datos con lógica paraconsistente para la seguridad en el metaverso.*

2.2 Estrutura do Trabalho

Em um ambiente organizacional podem ser detectados muitos possíveis eventos de incidentes envolvendo quebra de privacidade dados diariamente, sendo que um *Data Loss Prevention - DLP* Paraconsistente analisará quais são os reais incidentes de privacidade que devem ser acompanhados e tratados, descartando os falsos-positivos.

As fases do projeto estão definidas na figura 2. Durante o andamento do projeto estas fases passaram por três tipos de *status*: não efetuadas (cor cinza), em andamento (cor amarela) e efetivadas (cor verde). Sendo que todas elas já foram efetivadas e detalhadas no Capítulo 4.

Figura 2: Fases do Projeto:

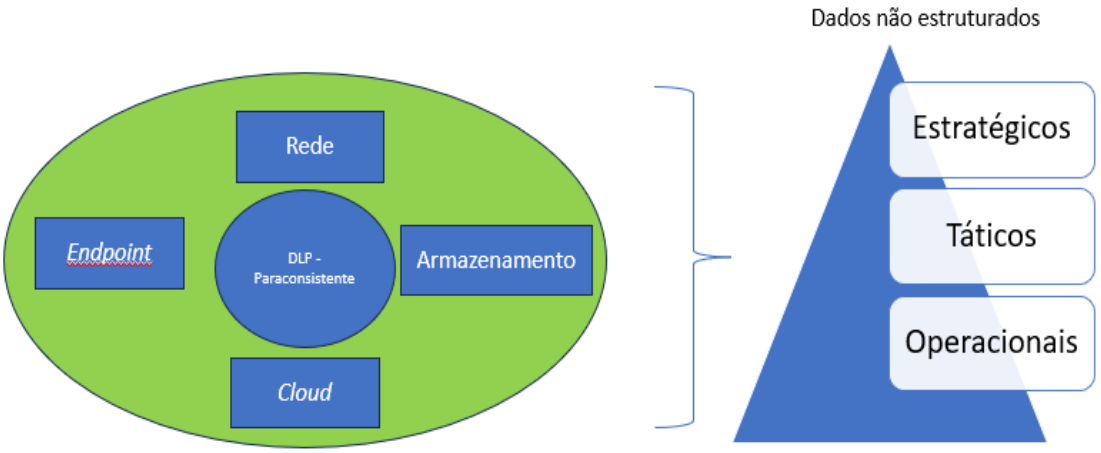


Fonte: Autora

O foco desse tipo de projeto visa a análise de dados não estruturados localizados nas corporações em várias áreas como: na rede, nos *endpoints* (que são os equipamentos utilizados pelos usuários, tais quais, computadores *desktops*, *notebooks*, *smartfones*, *tablets* e *smartwatch*), áreas de armazenamento *on premisses* (servidor local) ou *cloud* (em nuvem).

A ferramenta de captura e detecção desses dados para verificação da perda/dano/roubo/evasão, neste projeto é realizado pelo DLP Paraconsistente conforme explicação gráfica da Figura 3.

Figura 3: Estrutura da captura de dados pelo DLP Paraconsistente



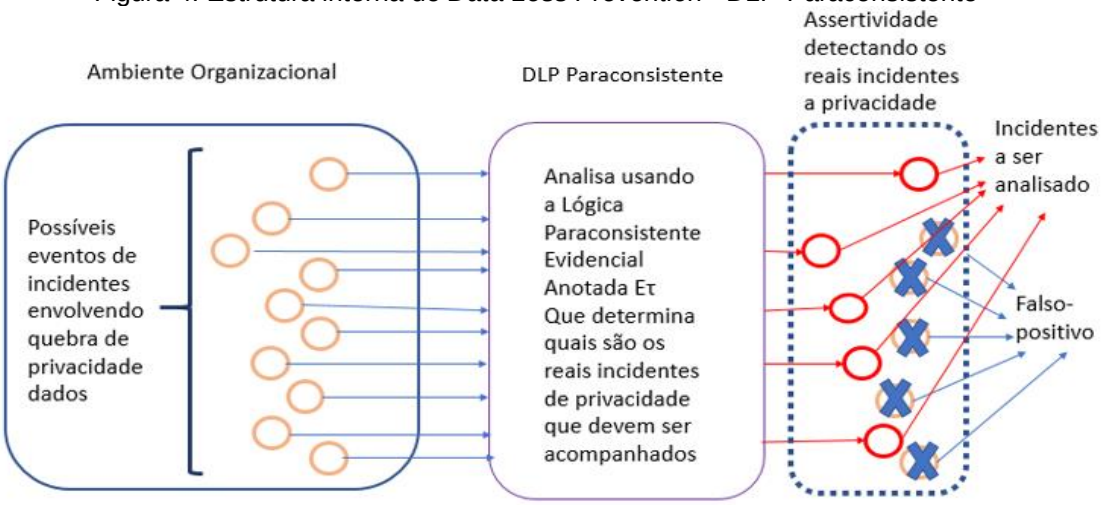
Fonte: Autora

Depois que os dados são capturados ainda é realizada uma pré-análise, ou seja, são limpos/anonimizados em conformidade com a Lei Geral de Proteção de Dados - LGPD e são organizados em níveis da hierarquia de atividades processuais da corporação, onde são segmentados em dados estratégicos, táticos e operacionais.

Dentro do Sistema de Gerenciamento de Dados Pessoais diariamente ocorrem várias detecções de eventos de possíveis incidentes sobre quebra de privacidade de dados. Estas situações podem chegar a quantidades exponenciais o que dificulta a análise de um encarregado de dados, mas nesse projeto haverá a concentração somente para os casos de dados não estruturados, como os que estão dentro de e-mails, mensagens de redes sociais (*Instagram, LinkedIn, Facebook*, entre outros), e de que forma são identificados com incidentes. Observando que o tempo de ação, ou seja, a resposta para minimização do incidente e correção ou remediação da situação deve ser urgente, mas devido as grandes quantidades de alertas torna-se inviável uma ação imediata sem uma pré-análise o que levaria a necessidade da contratação de centenas de pessoas somente para este tipo de ação. Muitas organizações contratam empresas terceirizadas para este tipo de atuação de monitoramento

de segurança da informação, que denominam como Centro de Operações de Segurança em Tecnologia da Operação ou *Security Operations Center - SOC*. Mas nem todas as corporações possuem orçamento e preferem correr o risco de ataques cibernéticos. Por isso, a estrutura de um *Data Loss Prevention - DLP* Paraconsistente pode auxiliar na assertividade dos incidentes, conforme figura 4.

Figura 4: Estrutura interna do *Data Loss Prevention - DLP* Paraconsistente



Fonte: Autora

CAPÍTULO 3 – REVISÃO BIBLIOGRÁFICA

3.1 Lógica Paraconsistente Anotada Evidencial Eτ

A Lógica Anotada Evidencial Eτ pertence ao grupo de lógicas não clássicas. Elas são um tipo de lógica paraconsistente e lógica paracompleta, ou seja, é uma lógica não alética. Uma lógica diz paraconsistente se puder ser a lógica subjacente de teorias inconsistentes, mas não triviais, ou seja, na lógica paraconsistente, há fórmulas e suas negações verdadeiras. Uma lógica diz paracompleta se puder ser a lógica subjacente de teorias em que há fórmulas P e ¬P (a negação de P), ambas falsas. Diz-se que uma lógica é não alética quando ela é ao mesmo tempo paracompleta e paraconsistente. A lógica Eτ é uma lógica não alética adequada para um raciocínio evidencial.

Neste trabalho, considera-se um tipo especial de lógica anotada, a Lógica Paraconsistente Anotada Evidencial Eτ - Lógica Eτ.

3.1.1 Estado da arte

ABE (1992) cita que as Lógicas Paraconsistentes Anotadas são uma família de lógicas não clássicas surgidas no início da década de 90 do século passado em programação lógica.

As lógicas anotadas estão relacionadas a certos reticulados completos, que desempenham um papel importante.

Em ABE (1992): “A Lógica Paraconsistente Anotada Evidencial Eτ possui uma linguagem e as proposições atômicas são do tipo p(μ, λ) onde p é uma proposição e μ, λ∈ [0, 1] (intervalo real unitário fechado). Intuitivamente, μ indica o grau de evidência[1]favorável de p e λ o grau de evidência contrária de p. A leitura dos valores μ, λ dependem das aplicações consideradas e podem sofrer mudanças: com efeito μ pode ser o grau de crença[2] favorável e λ poder ser o

---

1) O termo evidencia se encontra empregado num sentido não rigoroso, podendo intuitivamente ser “certeza” manifesta ou dados e informações que suportam opiniões. O termo “grau de evidência” significa o que se está explanado no curso do trabalho.

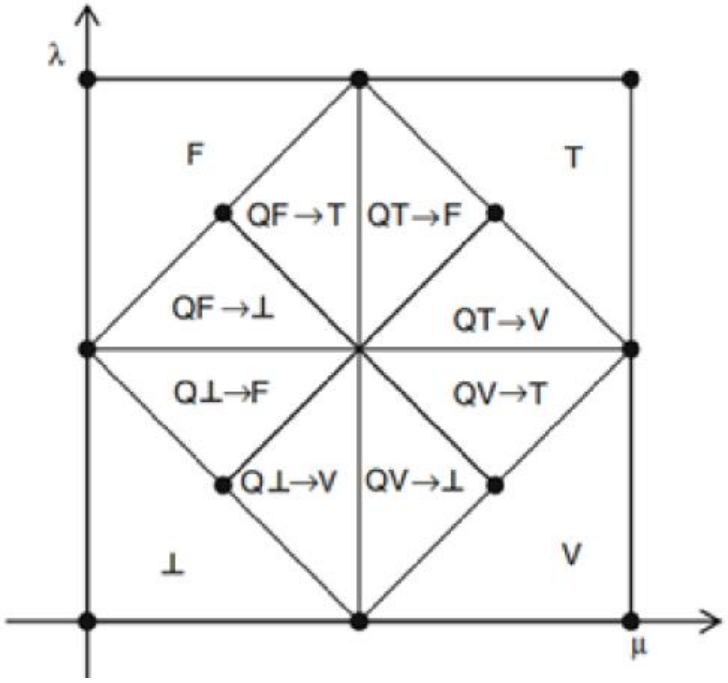
2) O termo crença também se encontra empregado em um sentido não rigoroso. Convém ressaltar que usualmente possui uma certa subjetividade.

grau de crença contrária da proposição p; também, μ pode indicar a probabilidade[3] expressa por p ocorrer e λ a improbabilidade expressa por p de ocorrer. As proposições atômicas p(μ, λ) da lógica Eτ podem ser intuitivamente ser lidas como: creio em p com o grau de crença favorável μ e o grau de crença contrária λ, ou o grau de evidência favorável de p é μ e o grau de evidência contrária de p é λ".

Conforme resumido de ABE (1992): atualmente podem-se construir programas utilizando as lógicas paraconsistentes sendo possível o tratamento das inconsistências de um modo direto e elegante. Com esse recurso, pode-se aplicar em sistemas especialistas, banco de dados orientados a objetos, representação de conhecimento contraditório etc. com todas as implicações em Inteligência Artificial.

A representação da Figura 5 mostra uma representação do reticulado construído com valores de Graus de Certeza e de Contradição e seccionado em 12 regiões. Desse modo, no final da análise se obterá como resposta para tomada de decisão um dos 12 possíveis estados lógicos resultantes.

Figura 5: Representação do reticulado com os 12 estados: extremos e não extremos



Fonte: ABE (2015).

[3)]Atente-se que há diversas teorias de probabilidades.

A Tabela 1 a seguir apresenta o detalhe dos Estados Extremos e Não Extremos:

Tabela 1: Estados Extremos e Não Extremos e seus símbolos.

Estados Extremos	**Símbolo**	**Estados Não Extremos**	**Símbolo**
Verdadeiro	V	Quase Verdadeiro tendendo ao Inconsistente	$QV \rightarrow T$
Falso	F	Quase Verdadeiro tendendo ao Paracompleto	$QV \rightarrow \perp$
Inconsistente	T	Quase Falso tendendo ao Inconsistente	$QF \rightarrow T$
Paracompleto	$\perp$	Quase Falso tendendo ao Paracompleto	$QF \rightarrow \perp$
		Quase Inconsistente tendendo ao Verdadeiro	$QI \rightarrow V$
		Quase Inconsistente tendendo ao Falso	$QI \rightarrow F$
		Quase Paracompleto tendendo ao Verdadeiro	$Q\perp \rightarrow V$
		Quase Paracompleto tendendo ao Falso	$Q\perp \rightarrow F$

Fonte: ABE (2015).

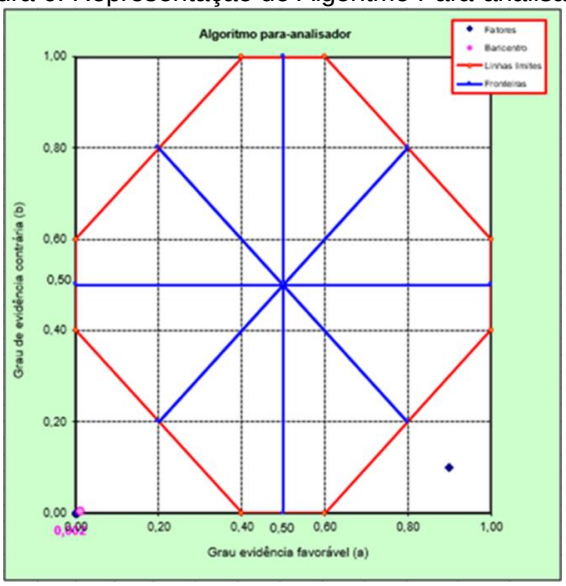
Com os cálculos dos valores dos eixos que compõem o reticulado pode-se reparti-lo ou delimitar internamente várias regiões, obtendo-se assim uma discretização dele. A partir das regiões delimitadas do reticulado, pode-se relacionar estados lógicos resultantes, os quais, por sua vez, serão obtidos pelas interpolações dos Graus de Certeza G_c e de Contradição G_{ct}. Dessa forma, para cada ponto de interpolação entre os Graus de Certeza e de Contradição haverá uma única região delimitada que o reticulado vai ser repartido depende da precisão pretendida na análise (DA SILVA FILHO et al., 2010).

Um especialista do conhecimento na temática a ser tratada emite sua opinião quantitativa que varia de 0,0 até 1,0. Esses valores são respectivamente a evidência favorável que é expressa pelo símbolo μ e a evidência contraria por λ.

3.1.2 Algoritmo Para-analisador

O algoritmo Para-analisador foi proposto na DA SILVA FILHO (1999), conforme Figura 6. Observa-se um conjunto de informações obtidas, que as vezes podem parecer contraditórias, dificultando a análise do cenário para análise do risco. Geralmente em tais situações estas informações são descartadas ou ignoradas, ou seja, são consideradas ruídos do sistema, porém na melhor das hipóteses podem até receber tratamento diferenciado. Exemplo da representação do Para-analisador pela Figura 6:

Figura 6: Representação do Algoritmo Para-analisador

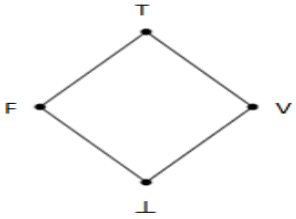


Fonte: DA SILVA FILHO (1999).

Nesta linha de raciocínio para a análise baseada na Lógica Paraconsistente serão consideradas situações de Inconsistência e Paracompleteza em conjunto com as Verdadeiras e Falsas.

O conjunto destes estados ou objetos ($\tau = \{F, V, T, \bot\}$) podem também ser chamados de constantes de anotação e podem ser representados por meio do diagrama de Hasse conforme mostra a Figura 7:

Figura 7: Diagrama de Hasse.



Fonte: (DA SILVA FILHO, ABE e TORRES, 2010).

"O operador sobre τ é: ~: $|\tau| \rightarrow |\tau|$ que operará, intuitivamente, assim:

~T = T (a 'negação' de uma proposição inconsistente é inconsistente)

~V = F (a 'negação' de uma proposição verdadeira é falsa)

~F = V (a 'negação' de uma proposição falsa é verdadeira)

$\sim\perp = \perp$ (a 'negação' de uma proposição paracompleta é paracompleta)

Será utilizada a Lógica Paraconsistente Anotada, este tipo deve ser composto por 1, 2 ou "n" valores.

3.2 Proteção e Privacidade de Dados

3.2.1 Estado da Arte

O Brasil instituiu LGPD – Lei Geral de Proteção de Dados com base na lei *GDPR – General Data Protection Regulation* da União Europeia. Porque isso foi importante, visto que as empresas Brasileiras que trocam informações com outros países não poderiam mais estar comercializando ou trocando informações sem o alinhamento quanto a privacidade de dados, situação que no mundo já era ponto primordial instituiu-se assim a LGPD – Lei 13.709/2018 (BRASIL, 2022).

Conforme o artigo 2º da LGPD entende-se a importância da proteção de dados e seus fundamentos, pois todas as pessoas possuem direito à privacidade, que é algo que vai além da segurança da informação, mas envolve o modo de viver das pessoas, sua individualidade, devido a envolver os direitos humanos, da liberdade de personalidade, da dignidade e da cidadania, entretanto indivíduos mal intencionados, os crackers querem se aproveitar de circunstâncias vulneráveis para conseguir tirar proveito desses direitos, e o papel do hacker ético é conseguir enxergar e reportar estas vulnerabilidades antes que elas sejam exploradas (BRASIL, 2022).

Com a entrada da LGPD o hacker ético também precisou se adequar à nova legislação, quanto sua atuação, com necessidade expressa de contratos com o controlador, que por sua vez solicitará o consentimento dos titulares de dados formalizado para manipulação de seus dados pessoais, e alguns

sensíveis. São considerados dados sensíveis: informações médicas, de orientação sexual, prática religiosa, a filiação a sindicatos, informações de menores de idade e adolescentes, entre outros.

Conforme PINHEIRO (2020) existem 11 princípios que a LGPD apresenta para o tratamento de dados pessoais, que são:

- A finalidade
  - Atividade de processar para propósitos legítimos, específicos, explícitos e informados ao titular.
- A adequação
  - Compatibilidade do tratamento com as finalidades informadas ao titular.
- A necessidade
  - Limitação do tratamento ao mínimo necessário para a realização de suas finalidades.
- O livre acesso
  - Garantia, aos titulares de consulta facilitada e gratuita sobre a forma do tratamento.
- A qualidade dos dados
  - Garantia, aos titulares, de exatidão, clareza, relevância e atualização dos dados, de acordo com a necessidade e para o cumprimento da finalidade de seu tratamento.
- A transparência
  - Refere-se aos titulares que precisam saber sobre todas as atividades de tratamento.
- A segurança
  - utilização de medidas técnicas e administrativas aptas e proteger os dados pessoais.
- A prevenção
  - Medidas adotadas para prevenção de ocorrência de danos em virtude do tratamento de dados pessoais.
- A não discriminação
  - Ausência de possibilidade na realização do tratamento de dados para fins discriminatórios ilícitos ou abusivos.

- A responsabilização
    - Demonstra que o agente adota medidas eficazes e capazes de comprovar a conformidade.
- A prestação de contas
    - Prestação de contas pelo agente, da adoção de medidas capazes de comprovar a proteção de dados pessoais.

3.2.2 Componentes da LGPD

Os principais atores dentro desse contexto da privacidade de dados são: o Titular de dados, o Controlador, o Operador, o encarregado de dados ou *DPO – Data Protection Officer*, a ANPD – Autoridade Nacional da Privacidade de Dados (ANPD, 2021). Sendo que desses são considerados agentes de tratamento somente o Controlador e o Operador, o encarregado de dados ou *DPO – Data Protection Officer* pode ser contratado pessoa física ou pessoa jurídica por qualquer um deles (BRASIL, 2022).

3.2.3 O Titular de dados

O titular dos dados é a pessoa física que concorda em compartilhar suas informações com algum controlador, isto é, quando uma pessoa faz um cadastro em um site para comprar um sapato, roupa ou objeto, neste momento expõe suas informações pessoais a uma empresa que irá manipular suas informações pessoais, como: nome, endereço, CPF, RG, dados de cartão de crédito, entre outros. Diante do artigo 18 da LGPD (BRASIL, 2022) o titular de dados possui alguns direitos, que são:

- Confirmação de que existe tratamento;
- Possibilidade de acesso aos seus dados pessoais;
- Possibilidade de correção de seus dados quando estão incompletos, inexatos ou desatualizados;
- Possibilidade de solicitar anonimização, bloqueio ou eliminação de dados que já não são mais necessários;
- Possibilidade de se fazer a portabilidade de dados;
- Quando não existir mais necessidade por parte do controlador, poder consentir na eliminação os seus dados;

- Poder solicitar informações sobre os dados pessoais que foram compartilhados com o controlador;
- Entender sobre a consequência do não consentimento dos dados pessoais com o controlador;
- Poder solicitar a qualquer momento a revogação do consentimento aos seus dados pessoais.

3.2.4 O Controlador e o Operador

O Controlador pelo artigo 5 da LGPD (BRASIL, 2022) é quem faz a tomada de decisões quanto ao tratamento de dados pessoais dos titulares sob a sua responsabilidade, e o Operador é aquele que efetiva propriamente o tratamento desses dados no lugar do Controlador, ou seja, é o que realmente pratica o processo de tratamento dos dados. Sendo que o Operador deve ser sempre uma pessoa distinta do Controlador, nem mesmo seu subordinado ou empregado de alguma de suas empresas ou do grupo associado (ANPD, 2021).

Outro ponto importante que se deve ressaltar é que o sistema do Controlador deve delimitar a tomada de decisão quanto a finalidade do uso dos dados, bem como do seu tempo de uso explícito ao Titular de dados, vulnerabilidade essa explorada por muitos *Crackers* para sequestro de dados indevidamente coletados.

Na União Europeia foi promulgada a Regulamento geral de Proteção de Dados - GDPR para normatizar com uma visão mais acurada pontos de tomada decisão até conjunta quando existem dois Controladores atuando em conjunto, situação que pode existir em um contrato mútuo de prestação de serviço para um hacker ético (*EUROPEAN COMISSION*, 2016).

Para a LGPD (BRASIL, 2022) existem 10 tipos bases legais que um Controlador ou Operador podem trabalhar e são estes pontos que devem ser explorados por *hacker*, que são:

- Informações consentidas pelos Titulares de Dados
    - Bases dos consentimentos dos titulares de dados, ou seja, é uma base de dados com as declarações que os titulares concordam com o uso de seus dados pessoais pela empresa para uma determinada finalidade.

- Dados para cumprimento de alguma obrigação legal
    - Nesta situação, o tratamento de dados se justifica pela obrigação de cumprir alguma legislação.
- Informações para execução de políticas públicas
    - Quando se aborda a administração pública, os dados podem ser tratados para a execução de políticas públicas previstas em leis, contratos, convênios ou similares.
- Estudos realizados por Órgão de Pesquisa
    - Neste caso, o tratamento de dados é válido, pois as instituições públicas e privadas podem fazer estudos e pesquisas com intuito de desenvolvimento científico, social ou econômico.
- Informações para Contrato ou diligências pré-contratuais
    - Quando duas partes formalizam um contrato com termos que permitem o uso de dados pessoais, o tratamento de dados pode ser realizado.
- Dados para o exercício regular dos direitos dos titulares
    - Conforme o art. 7º. – LGPD, o tratamento de dados pode ser feito para o exercício regular de direitos em processo judicial, administrativo ou arbitral – Lei de Arbitragem – Lei nº 9.307/1996 (BRASIL, 1996).
- Informações necessárias para proteção da vida
    - Justifica-se o uso de dados pessoais quando é indispensável para a proteção da vida e da segurança física do titular, sem precisar de seu consentimento. Por exemplo, em caso de um acidente e o titular não poder se comunicar e seus dados serem necessários para uma internação hospitalar.
- Tudo que tem a ver com a tutela da saúde
    - Disponibilização de dados para profissionais de saúde, serviços de saúde ou autoridades sanitárias para tratamento de dados com finalidade válida, como: prontuários, Classificação Internacional de Doenças e Problemas Relacionados com a Saúde - CID, tratamentos, internações, tipo sanguíneo, comorbidades etc.
- Informações sobre interesses legítimos do Controlador/Operador

    - Refere-se ao apoio e promoção de atividades do controlador e a proteção, em relação ao titular, do exercício regular de seus direitos ou prestação de serviços que o beneficiem, respeitadas as legítimas expectativas dele e os direitos e liberdades fundamentos, conforme o art. 7º, inciso IX e o art. 10 da LGPD (BRASIL, 2022).
- Informações referentes a proteção ao crédito
    - Os dados pessoais podem ser consultados para avaliar o histórico de crédito da pessoa, ou seja, para a aprovação de crédito e a redução de riscos de transação.

3.2.5 O Encarregado de Dados

O encarregado de dados ou *Data Protection Officer – DPO* não é especificamente um agente de tratamento de dados, mas uma figura delineada e obrigatória pela LGPD para fazer uma ponte com o Titular de Dados em seus questionamentos com o Controlador/Operador, e quando a situação envolve uma informação direta a Autoridade Nacional de Proteção de Dados – ANPD quando da ocorrência de algum incidente de segurança (artigo 48 da LGPD, BRASIL, 2022). A atividade de Encarregados de Dados é regulamentada pelo Ministério do Trabalho especificamente pela CBO – Classificação Brasileira de Ocupações – 1421-35, onde é denominado como Oficial de proteção de dados pessoais – dpo (CBO, 2023).

Ele deverá coordenar e levar as empresas à conformidade, não só da LGPD (BRASIL, 2022), mas de outras normas, como: Marco Civil da Internet, ISO 27.701, ISO 27.001, que serão aplicadas nos seus ambientes.

Além de mapear os processos: do ciclo de vida da informação, da responsabilidade das pessoas envolvidas e das tecnologias utilizadas para suporte (BRASIL, 2022).

Observando-se que no caso da *Cambridge Analytica* são apresentados dados de titulares de outros países além dos Estados Unidos. Assim pode-se alertar sobre a importância da ética e da privacidade dos dados e como um profissional de segurança da informação em conjunto com o encarregado de dados podem auxiliar no aprimoramento dos controles e minimizar as vulnerabilidades dos sistemas com suas ferramentas e habilidades em detectar

lacunas facilmente. Este é um exemplo negativo, conforme tabela 2, onde ocorreram o vazamento de dados de 87 milhões de usuários do *Facebook* de forma indevida, sendo que os dados foram manipulados sem o consentimento dos seus titulares, com intuito de ajudar na campanha presidencial dos Estados Unidos a favor do ex-presidente americano *Donald Trump* (HINDS, 2020):

Tabela 2: *Cambridge Analytica*

Caso *Cambridge Analytica* Países prejudicados	**Quantidade de Dados Vazados do Facebook pela *Cambridge Analytica***
Estados Unidos	70.632.350
Filipinas	1.175.870
Reino Unido	1.079.031
México	786.880
Canadá	622.161
Índia	562.455
Brasil	443.117
Vietnã	427.446
Austrália	311.127
Indonésia	109.666

Fonte: HINDS (2020).

3.3 *DLP – Data Loss Prevention*

3.3.1 Estado da Arte

Data Loss Prevention não é novidade, mas está sendo cada vez mais utilizado, pois ferramentas prontas no mercado já possuem alguns recursos deste padrão de monitoramento embutidos. Aplicativos como o *Office 365* já conseguem trazer alguns *templates* padronizados internamente para conformidade com o Regulamento Geral de Proteção de Dados - GDPR (*EUROPEAN COMMISSION*, 2016). Mesmo tentando fazer uma análise para o cumprimento das normas, com a falta de profundidade na avaliação de determinados critérios, pode ocorrer perda de dados (SILOWASH, 2013).

Outras ferramentas de *Antivírus* como o *Trend Micro* (LIU, 2007), que possui a solução chamada *Deep Security* especificamente para servidores, e

outra solução chamada *Apex One*, que é um *Software as a Service* - SaaS para segurança de *endpoint dos usuários*, ambas ferramentas possuem a funcionalidade de reportar *logs* DLP. Estes logs podem ser configurados de acordo com as necessidades da empresa, por exemplo: busca de informações de divulgação de CPF, CNPJ, Notas Fiscais, cartão de crédito, previdência etc. (SILOWASH, 2013).

Outra ferramenta é *Purview* da *Microsoft* que auxilia da mesma forma configurando-se a detecção de logs com busca em informações ou palavras chaves, podendo ser configurado para analisar o conteúdo transitado de *e-mails* (AHMAD, 2023).

Internamente, cada uma dessas ferramentas trabalha com algoritmos de Inteligência Artificial que cruzam dados e monitoram padrões pré-formatados pelo analista de segurança da informação, mas usando diretamente a lógica clássica para suas conclusões e apresentações de relatórios para que um gestor possa tomar uma decisão. Entretanto a quantidade de dados gerados é bem vasta, criando a necessidade de muitas pessoas para análise minuciosa de cada tipo de log designado como importante, chegando a milhares de dados selecionados para uma busca, por exemplo, CPF apresentado dentro de e-mails, necessitam de consentimento dos titulares para que esta informação transite entre funcionários, áreas da empresa e muito mais quando são enviados para fora do ambiente interno. Nesta situação já existe um incidente de privacidade de dados, se não houver consentimento explícito do titular para essa situação (MALDONADO e BLUM, 2022).

A Decisão de continuar com o processo quando uma situação de perda de dados não estruturados, como os que estão dentro de e-mails é identificada como um ponto de possíveis incidentes e deve ser analisada detalhadamente (MALDONADO e BLUM, 2022).

Data Loss Prevention - DLP ajuda a delimitar os dados:

• Quais são suas origens, usuários, áreas ou insumos?

• Onde estão armazenados?

• Dentro de quais sistemas os dados passam para consistência e se transformam em relatórios informativos?

• Quais são as saídas de dados?

• Quando, como e onde podem ser destruídos?

Pode-se delimitar a fonte de dados estruturados, localizados principalmente em:

• Da Rede (nos servidores e sistemas de aplicação);

• Da nuvem (armazenamento externo);

• Do armazenamento (armazenamento local);

• Dos *endpoints* (armazenamento do usuário final).

Enquanto os não estruturados estão geralmente nos:

• Dados pessoais que transitam dentro dos e-mails dos usuários de uma empresa;

•Relatórios impressos, como relatórios antigos de sistemas legados desativados;

• Dados armazenados em microfilme ou microficha, tecnologia descontinuada, mas muitas vezes armazenadas devido ao valor jurídico;

• Dados em redes sociais como *WhatsApp, Instagram* e *Facebook*;

• Dados em pen drives e discos rígidos externos;

• Dados em unidades de nuvem, como *OneDrive, SharePoint, Teams, Google drive etc.*

O DLP precisa de uma configuração assertiva por um analista especialista na área para capturar possíveis resultados com as necessidades da empresa, como padrões de compra de cartão de crédito, que foi uma das áreas que primeiro desenvolveu padrões para monitoramento de situações suspeitas que podem levar a perda/vazamento/roubo de dados (SIKORSKI, 2012).

Outra situação amplamente utilizada para padronização de DLP é para suspeita de *malware*, *ransomware* e spam, pois alguns indicadores podem ser copiados de Listas de Segurança Pública chamadas *Common Vulnerabilities and Exposures* - CVEs. Estas listas de segurança são geralmente incluídas dentro das ferramentas padrões de mercado e atualizadas conforme são emitidas pelos órgãos de segurança renomados como o *Mitre framework* (RAJESH,2022).

Nem todas as ferramentas de detecção de incidentes podem oferecer a seus clientes serviços de otimização muito competitivos; no entanto, muitos precisam de um gerenciamento muito refinado e cuidadosamente protegido, geralmente por um *Security Information and Event Monitoring* - SIEM, ou seja, uma ferramenta que consegue unificar alertas sobre atividades maliciosas como varredura de IP, fluxo de dados, e-mails maliciosos, intrusão, tentativas de *spam*, monitoramento de *firewall*, revisão de atualizações de política de segurança,

patches de segurança implementados, mas aqui usaremos esse conceito para criar um processo específico focado em perda/dano/roubo/vazamento de dados (SIKORSKI, 2012).

Além disso, melhores oportunidades para identificar efetivamente elementos redundantes e não ideais em estruturas organizacionais e de gestão que podem ser identificados, alterados ou removidos com segurança, trazendo o benefício da precedência compra da ferramenta de mercado antes da definição da estrutura de gerenciamento de dados (SILOWASH, 2013).

Se houver demanda em relação ao ambiente de negócios desatualizado, conscientiza-se sobre a necessidade de realizar uma otimização completa da estrutura de gerenciamento de segurança da organização e privacidade de dados utilizando preferencialmente aspectos de DLP para evitar perda ou vazamento de dados (SINGH,2018).

As ferramentas geralmente precisam de algum tempo de teste e calibração, para que não criem alertas com excesso de falsos positivos, mas usem a lógica clássica (MORIN et al, 2009).

O nível de sofisticação alcançado pelo software malicioso exige esforços constantes e considerável dispêndio de recursos para mitigar essa prática. O equilíbrio do apetite ao risco entre as partes é frágil, conforme reconhecido por diversas autoridades da área. Portanto, é preciso inovar constantemente e estar um passo à frente desses ofensores (SINGH,2018).

Ao utilizar o conceito de DLP, a seguinte granularidade foi considerada para obter a graduação para os tipos de perda de dados detectados (SILOWASH, 2013):

Ferramentas de DLP no mercado tem sido cada vez mais buscadas pelas empresas devido a LGPD no Brasil, GDPR (EUROPEAN COMMITE, 2022) na União Européia. O refinamento do controle interno de acesso aos dispositivos deve ser granular referente a necessidade de acesso, como uso de dispositivo USB, tablets, smarthwatches, entre outros.

O principal objetivo é sempre de bloqueios, monitoramento e gerenciamento de dispositivos. O controle granular com base na identificação (ID) do distribuidor, fornecedor, cliente, identificação do produto (ID), número de série são explorados pelos sistemas (MORIN et al, 2009).

Há diversos canais que podem ser utilizados em dados que passam pela digitalização do movimento, sendo plausível de monitoramento e bloqueio nas transferências de arquivos. Todo o detalhe é feito por conteúdo e contexto.

Constantemente são fornecidos os dispositivos de manipulação ou automáticos com o intuito de fornecer dados e fornecer o DLP. Há casos que durante os dados em trânsitos são criptografados de maneira forçada e assim mantêm a qualidade da técnica DLP no tratamento de dados (SCHALLER, 2022).

Pela LGPD entende-se que o tratamento deve ser considerado como operação realizada com dados pessoais, como se refere a coleta, produção, recepção, classificação, utilização, acesso, reprodução, transmissão, processamento, arquivamento, armazenamento, eliminação, ou controle da informação, transmissão, comunicação, envio, difusão ou avaliação (BRASIL,2022).

Atualmente, foi instituída pela Autoridade Nacional de Proteção de Dados - ANPD a delimitação de multas e até a paralisação das atividades da empresa em caso do nível de discordância da LGPD, conforme Regulamento de Dosimetria publicado em 27/02/2023 (ANPD, 2023).

Internacionalmente provenientes de fora do território nacional e não são objetos de comunicação, uso compartilhado de dados com agentes de tratamento brasileiros ou objeto de transferência internacional de dados com outro país que não seja de proveniência, que proporcione o país de proveniência grau de proteção de dados pessoais adequados ao previsto da LGPD (BRASIL, 2022).

O uso de ferramentas DLP como foco o atendimento da regulamentação a respeito de pessoal sensível: dados médicos, tipo racial ou dado étnico, convicção religiosa, opinião política, filiação sindical a origem ou a organização de caráter religioso, filosófico ou político, dado religioso ou a vida sexual referente, dado genético ou biométrico, quando vinculado a uma pessoa natural (BRASIL, 2022).

As empresas devem usar o conceito DLP pois tratam dados conforme os preceitos do operador e/ou controlador. Este é considerado a última pessoa natural jurídica, de direito público ou privado, a quem competem ou decisões judiciais ao tratamento de dados pessoais. E quando pessoa natural ou jurídica, de direito público ou privado, realiza o tratamento de dados pessoais em nome do controlador (KHAN,2021).

3.4 Ameaças Cibernéticas

3.4.1 Estado da Arte

No final, as empresas tendem a subestimar o nível de ameaças cibernéticas, o que abre boas oportunidades para implementação de um teste de penetração (tipo *Blackbox*) adequado do perímetro externo do negócio, para adequar os componentes de seu *cyber* ambiente (SHAH, 2015). Este tipo de solução não é recente, mas a sua popularização foi difundida após 2020.

Por isso, que as empresas começaram a levantar ativamente questões de reavaliação de vulnerabilidades e proteção dos perímetros externos e internos de infraestrutura crítica e empresas de vários setores. Pois apesar de medidas de segurança implantadas, não existe certeza de que a ameaça foi totalmente afastada (MENEZES, 2015).

O objetivo é otimizar uma parte organizacional e gerencial da infraestrutura para construir um sistema de segurança cibernética eficaz na privacidade de dados. O desenvolvimento do setor de Segurança de Privacidade de Dados tem demonstrado repetidamente mais complexidade e exigindo muita confiança dos gestores das tarefas operacionais para otimização desse tipo de monitoramento, principalmente pela escassez de profissionais especializados nesta área, bem como ferramentas de apoio a estes. Visto nenhum sistema ser 100% seguro como aborda o *framework NIST* (SHEN, 2014).

Desta forma, a proteção total bem-sucedida contra ameaças cibernéticas não necessita apenas do conhecimento técnico, mas alinhar o atendimento a grupo específicos, como diretores/executivos, gerentes operacionais, analistas de negócio e profissionais de recursos humanos e ferramentas de apoio. A

análise de risco nestes casos pode se apoiar em vários frameworks como, *NIST* (SHENS, 2014), *Mitre framework* (RAJESH, 2022 e STROM, 2018), *Top 10 OWASP* (BACH-NUTMAN, 2020 e MARCHAND-MELSOM, 2020), entre outros.

A estrutura do empreendimento organizacional e seus *cybers* ambientes por parte das empresas, e a complexidade dos métodos de ataques sofridos a cada dia cria a necessidade crescente de uma estrutura para monitoramento dos ambientes de forma segura e reforçada (PANDYA,2016).

Hoje em dia, muitas organizações estão buscando soluções de IA – Inteligência Artificial para encontrar analogias que levem a contenção otimizada podendo usar o foco também em lógicas não clássicas para contornar estes tipos de ataques de grupos mais ofensivos (CALVIN et al., 2023).

CAPÍTULO 4 – ELABORAÇÃO DO *DATA LOSS PREVENTION - DLP* PARACONSISTENTE

4.1 Sistema de Gerenciamento de Privacidade de Dados Pessoais

Nem todas as ferramentas de detecção de Incidentes conseguem oferecer aos seus clientes serviços de otimização bastante competitivos, entretanto muitas precisam de uma gestão muito refinada e cuidadosamente protegida por geralmente por um *Security Information and Event Monitoring - SIEM*, ou seja, uma ferramenta que consegue unificar os alertas sobre atividades maliciosas, como: verificação de IPs, fluxo de dados, e-mails maliciosos, tentativas de invasão, monitoramento de *firewalls*, análise de atualizações de políticas de segurança, patchs implementados, mas aqui usaremos este conceito para criar um processo específico voltado para o Gerenciamento de Incidentes de Privacidade de Dados (KHAN, 2021).

Além disso, melhores oportunidades para identificar com eficácia os redundantes e elementos não ideais na estrutura organizacional e de gerenciamento que podem ser identificados, alterados ou removidos com segurança, trazendo o benefício de uma prevenção antes de um incidente grave de segurança da privacidade de dados (BACH-NUTMAN, 2020).

Caso se perceba uma demanda quanto a desatualização do ambiente de seu negócio, traz a conscientização da necessidade de se realizar uma otimização completa da organização e estrutura de gestão da segurança da privacidade de dados (STROM, 2016).

As ferramentas geralmente precisam de um tempo de teste e calibração para que não crie alertas com falsos positivos excessivos, mas utilizam a lógica clássica (AKAMA, 2016).

4.2 Definições e Proposições Sobre Aspecto Regulatório

A LGPD (BRASIL, 2022) detalha em seu artigo 5º o ciclo de vida dos dados, que são determinados pelos seguintes passos:

- Criação que está determinado na: coleta, produção, recepção ou extração dos dados;

- Transporte por meio da: transmissão, distribuição e comunicação dos dados;

- Manuseio centrado na: classificação, utilização e modificação dos dados;

- Armazenamento determinado em: arquivamento e armazenamento;

- Descarte quando se dá a eliminação dos dados.

4.3 Escolha de especialistas

A seleção deve ser feita dentro dos 27 tipos de especialistas:

- Encarregado de dados, ou seja, *Data Protection Officer - DPO*;
- Analistas de Segurança da Informação;
- Analistas de Infraestrutura;
- Analistas de Monitoramento de Ambiente;
- Analista de *Telecom*;
- Desenvolvedor/Programador e;
- Analista de *Firewall*.

Seguem a seguir amostra das respostas do agrupamento dos especialistas para os fatores/seções:

Tabela 3: Agrupamento de especialistas do Grupo 1, 2 e 3

FATORES	**SEÇÕES**	**Grupo 1 - Encarregado de Dados ou DPOs**						**Grupo 2 - Analista de Segurança da Informação**						**Grupo 3 - Analista de Infraestrutura**					
		Expert 1		**Expert 2**		**Expert 3**		**Expert 4**		**Expert 5**		**Expert 6**		**Expert 7**		**Expert 8**		**Expert 9**	
		μ	**λ**	**μ**	**λ**	**μ**	**λ**	**μ**	**λ**	**μ**	**λ**	**μ**	**λ**	**μ**	**λ**	**μ**	**λ**	**μ**	**λ**
F1	**S1**	0,9	0,1	0,9	0,1	0,9	0,3	0,9	0,2	1,0	0,3	0,8	0,1	0,9	0,3	1,0	0,1	1,0	0,2
F2	**S1**	0,2	1,0	0,3	1,0	0,3	0,9	0,4	0,6	0,3	0,7	0,2	0,9	0,1	0,8	0,2	0,9	0,3	0,9
F2	**S2**	0,2	0,2	0,1	0,5	0,1	0,1	0,2	0,4	0,3	0,1	0,4	0,5	0,2	0,1	0,1	0,9	0,1	0,1
F2	**S3**	0,9	0,1	0,7	0,3	0,6	0,8	0,4	0,6	0,8	0,2	0,5	0,6	1,0	0,7	0,6	0,5	0,7	0,5

Fonte: Autora

Tabela 4: Agrupamento de especialistas do Grupo 4, 5 e 6

FATORES	SEÇÕES	Grupo 4 - Analista de Monitoramento de Ambiente						Grupo 5 - Analista de Telecom						Grupo 6 - Desenvolvedor					
		Expert 10		Expert 11		Expert 12		Expert 13		Expert 14		Expert 15		Expert 16		Expert 17		Expert 18	
		μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ
F1	**S1**	0,9	0,2	1,0	0,3	0,8	0,1	0,9	0,1	1,0	0,1	1,0	0,3	0,8	0,1	0,7	0,1	1,0	0,1
F2	**S1**	0,4	0,6	0,3	0,7	0,2	0,9	0,3	0,8	0,2	0,7	0,1	0,9	0,2	0,8	0,2	0,9	0,2	1,0
F2	**S2**	0,2	0,4	0,3	0,1	0,4	0,5	0,4	0,6	0,3	0,7	0,2	0,9	0,1	0,8	0,2	0,9	0,3	0,9
F2	**S3**	0,4	0,6	0,8	0,2	0,5	0,6	0,9	0,1	1,0	0,1	1,0	0,3	0,8	0,1	0,7	0,1	1,0	0,1

Fonte: Autora

Tabela 5: Agrupamento de especialistas do Grupo 7, 8 e 9

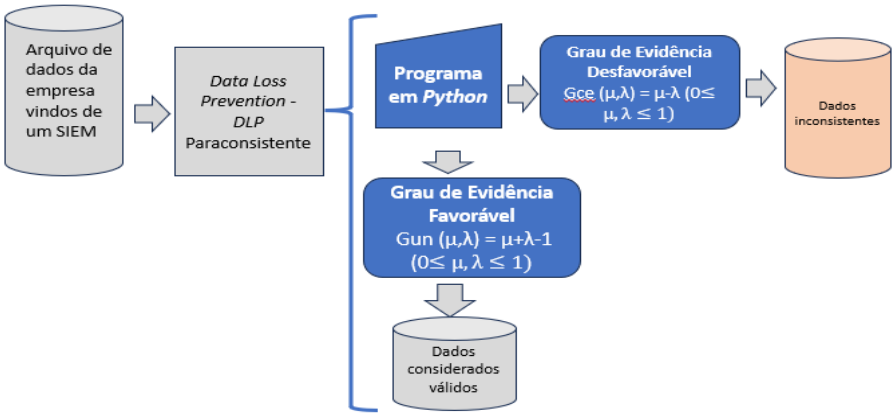
FATORES	SEÇÕES	Grupo 7 - Analista de SOC						Grupo 8 - Programador						Grupo 9 - Analista de Firewall					
		Expert 19		Expert 20		Expert 21		Expert 22		Expert 23		Expert 24		Expert 25		Expert 26		Expert 27	
		μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ
F1	**S1**	0,7	0,1	0,9	0,3	0,8	0,1	0,9	0,3	1,0	0,1	1,0	0,2	0,8	0,1	0,7	0,1	1,0	0,1
F2	**S1**	0,1	1,0	0,1	0,9	0,1	0,9	0,1	0,8	0,2	0,9	0,3	0,9	0,2	0,8	0,2	0,9	0,2	1,0
F2	**S2**	0,2	1,0	0,3	1,0	0,3	0,9	0,2	0,1	0,1	0,9	0,1	0,1	0,1	0,8	0,2	0,9	0,3	0,9
F2	**S3**	0,7	0,1	0,9	0,3	0,8	0,1	1,0	0,7	0,6	0,5	0,7	0,5	0,8	0,1	0,7	0,1	1,0	0,1

Fonte: Autora

4.4 Fixação de valores de controle

Neste estudo a fixação de valores de controle baseiam-se nos Graus de Evidência Favorável e Desfavorável, obtidos após os dados anonimizados serem processados pelo programa em *Python*, conforme figura 8:

Figura 8: Fixação de valores de controle



Fonte: Autora

4.5 Escolha de fatores de influência

Para o estudo de fatores de influência quanto aos incidentes de SIEM, conforme tabela 3:

Tabela 6: Fatores e Seções

Fatores	**Seções**
F1 – Análise de Vulnerabilidade	S1 - Comunicação com Rede *TOR*
F2 – Análise de Incidente	S1 - Suspeita de ataque de assinatura - tipo “*share*”
	S2 - Suspeita de ransomware – tipo “*create*” (O'KANE, 2018)
	S3 - Suspeita de Incidente de Privacidade de Dados (BRASIL, 2022)

Fonte: Autora

Uma rede *TOR* é uma rede anônima que proporciona potencial destrutivo e quebra a privacidade na Internet. Ao usar uma rede *TOR* um usuário pode ter liberdade para praticar crimes e serviços ilegais, para utilização da *Dark Web* (JANSEN, 2012).

Quando se fala de ataque de assinatura do tipo “*share*” é uma forma de invasão que pode acontecer quando um *hacker* tenta deixar algumas ferramentas dentro de um ambiente para poder utilizar para usar depois, com sua “assinatura” (DONG, 2019).

Já um *ransomware* é utilizado por *hackers* que automatizam seus ataques em empresas ou até com dados de pessoas físicas, onde são elaboradas as seguintes fases para concretização de extração de informações (O'KANE, 2018):

- Infecção da rede que irá ser atingida, nesta etapa usa um *ransomware* tipo “*share*”;
- Sequestro de dados com a encriptação, nesta etapa usa-se um *ransomware* tipo “*create*”;
- Solicitação de resgate de pagamento para devolução dos dados encriptados;
- Ativos ou informações devolvidas após o pagamento.

Enquanto, quando ocorre uma suspeita de Incidente de privacidade de dados, pode ser uma mescla das situações anteriores de risco cibernético, além de vazamento de alguma credencial pessoal por parte de um *hacker*. Esta situação não impacta somente a empresa, mas devido a regulamentação da LGPD (BRASIL, 2022), o titular envolvido do qual foi extraída a credencial ou dados pessoais deve ser informado e a Agência Nacional de Proteção de Dados - ANPD, reguladora governamental também deve ser informada em 72 horas com todos os parâmetros necessários e descritos na legislação, com possibilidade de punição para empresa caso a informação não ocorra, desde uma multa até a possibilidade de paralisação das atividades da empresa (ANPD, 2023).

4.6 Construção da Estrutura do *Data Loss Prevention - DLP* Paraconsistente

Programa em *Python* (Figura 9) e resultados da massa de dados.

Figura 9: Programa em *Python*:

```
import csv
from typing import List
from util import normalize_by_feature_scaling
from network import Network
from random import shuffle

if __name__ == "__main__":
    test_parameters: List[List[float]] = []
    test_classifications: List[List[float]] = []
    test_species: List[str] = []
    with open('\\Users\liliam\Downloads\RNTreinaCompara\RedeNeural_01\test.csv', mode='r') as test_file:
        tests: List = list(csv.reader(test_file))
        shuffle(tests) # get our lines of data in random order
        for test in tests:
            parameters: List[float] = [float(n) for n in test[0:4]]
            test_parameters.append(parameters)
            species: str = test[4]
            if species == "test_1":
                test_classifications.append([1.0, 0.0, 0.0])
            elif species == "test_2":
                test_classifications.append([0.0, 1.0, 0.0])
            else:
                test_classifications.append([0.0, 0.0, 1.0])
            test_3.append(species)

    normalize_by_feature_scaling(test_parameters)

    test_network: Network = Network([4, 6, 3], 0.3)

    def test_interpret_output(output: List[float]) -> str:
        if max(output) == output[0]:
            return "test_1"
        elif max(output) == output[1]:
            return "test_2"
        else:lor
            return "test_3"
```

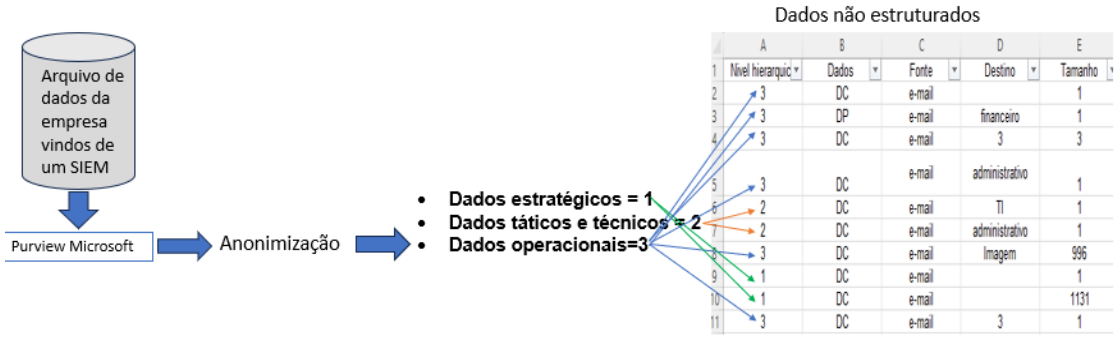
Fonte: Autora

A massa de dados a ser testada refere-se a uma empresa transportadora possuía o seguinte layout com esta legenda:

O primeiro nível de captação de dados reflete como eles podem ser reconhecidos pelo programa em Linguagem *Python*, conforme demonstra a figura 10:

- Dados estratégicos = 1
- Dados táticos e técnicos = 2
- Dados operacionais= 3

Figura 10: Anonimização dos dados recebidos



Fonte: Autora

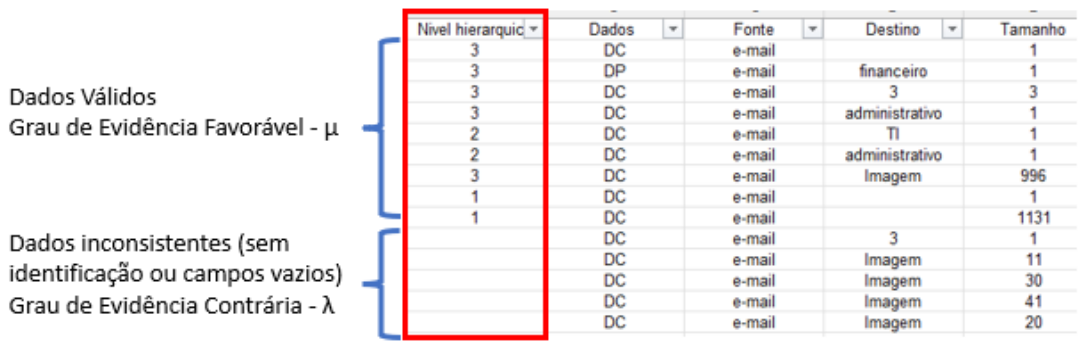
Quando os dados são extraídos do SIEM eles são gerados de forma bruta, ou seja, sem nenhum tratamento, entretanto para ser analisado sob o aspecto da LGPD ele deve ser anonimizado. Assim, cada e-mail identificado pelo SIEM como possível evento de incidente, passa pelo crivo de um analista que irá criar um arquivo somente com os itens necessários para análise do programa em Linguagem *Python*, que deve abrir o e-mail e identificar primeiramente se trata-se de uma informação:

- Estratégica;
- Tática ou técnica; ou
- Operacional.

Passando para o transporte dos dados para o arquivo em .CSV, em caso positivo para dado não estruturado de um e-mail com informações estratégicas é identificado na primeira coluna com o número 1. Depois se for identificado como dado tático ou técnico nessa primeira coluna é preenchido com o número 2. Mas quando é identificado como dado operacional, a primeira coluna é preenchida com o número 3. Estas identificações numéricas foram denominadas como Primeiro Nível de captação de dados.

Quando essa informação inicial da primeira coluna está preenchida corretamente, o dado ao ser analisado pelo programa em Linguagem *Python* é considerado como válido, onde se pode identificar o Grau de Evidência Favorável, ou seja, se possuir atribuição 1, 2 ou 3 na primeira coluna do arquivo. Caso não exista a possibilidade de identificação dessa atribuição e a primeira coluna apresentar zero ou campo vazio, a identificação é do Grau de Evidência Contrária, conforme figura 11:

Figura 11: Identificação do Grau de Evidência Favorável e Contrária nos dados do arquivo anonimizado – primeira coluna.



Nivel hierarquic	Dados	Fonte	Destino	Tamanho
3	DC	e-mail		1
3	DP	e-mail	financeiro	1
3	DC	e-mail	3	3
3	DC	e-mail	administrativo	1
2	DC	e-mail	TI	1
2	DC	e-mail	administrativo	1
3	DC	e-mail	Imagem	996
1	DC	e-mail		1
1	DC	e-mail		1131
	DC	e-mail	3	1
	DC	e-mail	Imagem	11
	DC	e-mail	Imagem	30
	DC	e-mail	Imagem	41
	DC	e-mail	Imagem	20

Fonte: Autora

Quando um *e-mail* é analisado e se verifica que se trata de informação pessoal no seu conteúdo a segunda coluna é preenchida com DP – Dados Pessoais, se o conteúdo abordar informações pessoais, mas o contexto é referente a empresa, com informações de sócios e/ou diretores, esta coluna é preenchida com DC – Dados Corporativos.

Quando essa informação da segunda coluna está preenchida corretamente, o dado ao ser analisado pelo programa em *Python* será considerado como válido, onde se pode identificar o Grau de Evidência Favorável, ou seja, se possuir atribuição DP ou DC na segunda coluna do arquivo. Caso não exista a possibilidade de identificação dessa atribuição e a segunda coluna apresentar zero ou campo vazio, a identificação será do Grau de Evidência Contrária, conforme figura 12:

Figura 12: Identificação do Grau de Evidência Favorável e Contrária nos dados do arquivo anonimizado – segunda coluna.

Nivel hierarquic	Dados	Fonte	Destino	Tamanho
3	DC	e-mail		1
3		e-mail	financeiro	1
3		e-mail	3	3
3		e-mail	administrativo	1
2	DC	e-mail	TI	1
2	DC	e-mail	administrativo	1
3	DC	e-mail	Imagem	996
1	DC	e-mail		1
1	DC	e-mail		1131

Fonte: Autora

As outras colunas do arquivo não influem diretamente na identificação do Grau de Evidência Favorável ou Contrária. A terceira coluna identifica o tipo de dado não estruturado que está sendo tratado é neste caso é único, ou seja, somente os dados de e-mails. A quarta coluna apresenta o destino ou destinatário, mas mesmo que não tenha identificação específica não altera a atribuição da identificação. Já a quinta coluna somente informa o tamanho do arquivo analisado, observando que a maioria não possui anexos e apresentam informações curtas.

Entretanto estes dados das colunas apesar de serem identificados conforme como dados válidos e inconsistentes, são dados brutos extraídos do SIEM e precisam ser normatizados para adequação ao programa em *Python* que utiliza visão da Lógica E_{τ}, conforme Tabela 7 para dados válidos:

Tabela 7: Normatização dos dados brutos do arquivo anonimizado para informações válidas

Localização no arquivo .CSV	Dado bruto	Dado normalizado
Atributo da coluna 1 - Nível de Hierarquia - Estratégica	1	1.0
Atributo da coluna 1 - Nível de Hierarquia - Tática	2	0.9
Atributo da coluna 1 - Nível de Hierarquia - Operacional	3	0.8
Atributo da coluna 2 - Dados Pessoais	DP	0.7
Atributo da coluna 2 - Dados Corporativos	DC	0.6

Fonte: Autora

Quando os dados brutos do arquivo normalizado apresentam informações em que constem “zero” ou ausência de informações, ou seja, o campo “vazio”, estas são consideradas inconsistentes, sendo a normalização apresentada na tabela 8:

Tabela 8: Normalização dos dados brutos do arquivo anonimizado para informações consideradas inconsistentes

Localização no arquivo .CSV	Dado bruto	Dado normalizado
Atributo da coluna 1 - Nível de Hierarquia - Estratégica	0 ou vazio	0.0
Atributo da coluna 1 - Nível de Hierarquia - Tática	0 ou vazio	0.1
Atributo da coluna 1 - Nível de Hierarquia - Operacional	0 ou vazio	0.2
Atributo da coluna 2 - Dados Pessoais	0 ou vazio	0.3
Atributo da coluna 2 - Dados Corporativos	0 ou vazio	0.4

Fonte: Autora

Segue um esquema para demonstrar como é feita esta identificação, conforme tabela 9:

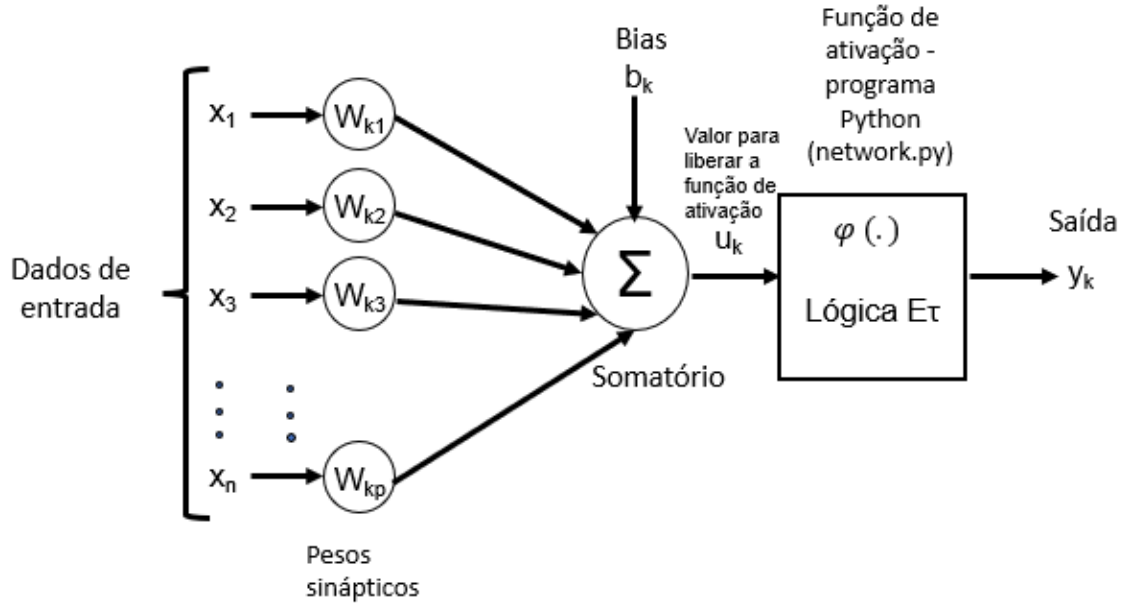
Tabela 9: Identificação na entrada de dados para o Grau de Evidência Favorável e Grau de Evidência Contrária

Localização no arquivo .CSV	Grau de Evidência Favorável	Grau de Evidência Contrária
Atributo da coluna 1 - Nível de Hierarquia - Estratégica	1.0	0.0
Atributo da coluna 1 - Nível de Hierarquia - Tática	0.9	0.1
Atributo da coluna 1 - Nível de Hierarquia - Operacional	0.8	0.2
Atributo da coluna 2 - Dados Pessoais	0.7	0.3
Atributo da coluna 2 - Dados Corporativos	0.6	0.4

Fonte: Autora

A análise de dados de entrada detalhada para o programa em *Python* foi baseada e adaptada do Neurônio Artificial de MCCULLOCH e PITTS (1943), onde os dados de entrada x_1 (dados estratégicos), x_2 (dados táticos e técnicos) e x_3 (dados operacionais), são os dados coletados e anonimizados, conforme definido no primeiro nível de captação de dados, que estão associados aos seus respectivos pesos (W_{k1} – peso 3 para x_1, W_{k2} – peso 2 para x_2, W_{k3}/W_{kp} – peso 1 para x_3). Sendo o peso 3 é o mais crítico, pois se referem a dados estratégicos, peso 2 para dados intermediários ou táticos e peso 1 para dados operacionais. O item denominado b_k – refere-se ao *Bias*, que representa a combinação linear de sinais de entrada seguido da função somatória (Σ), que permite um valor u_k que libera a função de ativação $\varphi(.)$. Nesta etapa no programa em Python (network.py) define os dados de saída da Lógica Eτ que apresenta o resultado de saída denominada y_k, ou seja, quantos desses dados são considerados válidos (Grau de Evidência Favorável) e quantos são considerados inconsistentes (Grau de Evidência Contrária), conforme Figura 13:

Figura 13: Adaptação do Neurônio Artificial de MCCULLOCH e PITTS



Fonte: MCCULLOCH e PITTS (1943).

Os Dados Estratégicos podem levar a perda de dados confidenciais de possíveis projetos ultrassecretos com maior risco ou com exposição da empresa no mercado, neste caso são somente referentes aos dados não estruturados dos *e-mails*.

Os dados táticos e técnicos podem ser representados pelos seguintes itens (HART, 2011):

- Identificação das áreas de negócio;
- Identificação de usuários impactados;
- Identificação do produto impactado;
- Identificação de incidentes com dados pessoais;
- Identificação do custo da multa por descumprimento da LGPD.

Caso ocorra vazamento de dados táticos e técnicos a situação pode levar a perda de informações confidenciais, como credenciais de: usuários chaves, de administradores de sistema, de administradores de banco de dados com autoridade para transferir grandes quantidade de dados ou de administradores financeiros que podem autorizar pagamentos de valores altos.

Os dados operacionais podem ser representados por:

- dados efetivamente identificados;
- dados contraditórios.

Em caso de perda de privacidade de dados operacionais, esta situação pode levar a perda de informações sobre padronização das atividades, gerando retrabalho ou necessidades iniciais de mapeamento, e novos treinamentos para as equipes de base. Em uma planta de produção industrial podem ser usuários operadores de maquinário especial e a subtração desse tipo de credenciais poderia acabar com uma linha de produção inteira ou até paralisar uma fábrica toda.

A massa de dados coletados da empresa financeira por apenas um mês é analisada pelo *Data Loss Prevention* - DLP Paraconsistente – Camada 1 – Programa em *Python* para obter o relatório com as respostas já pré-selecionadas quais são contraditórios.

Na camada 2 as funcionalidades podem ser utilizadas como entrada na Lógica E_T, pelo algoritmo Para-analisador:

- Verdadeiro;
- Falso;
- Incompleto;

- Paracompleto.

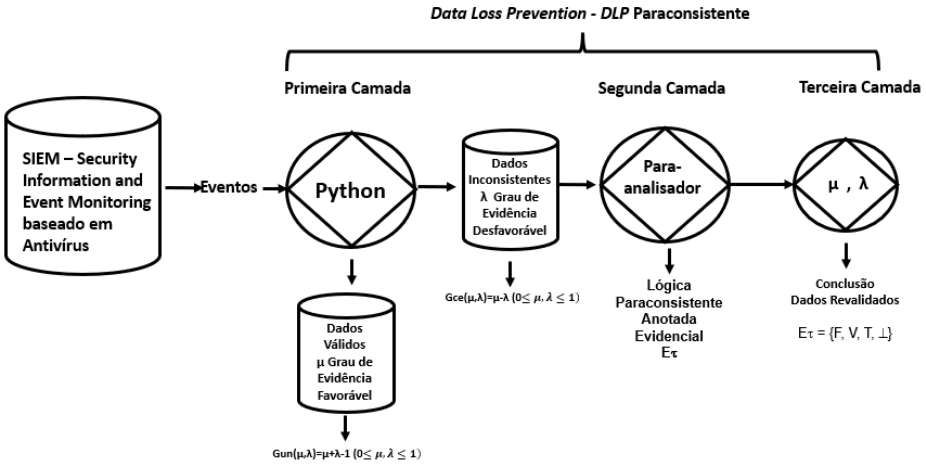
Na segunda camada (SILOWASH, 2013), as informações das exceções apresentadas pelo relatório do Algoritmo *Python* são reavaliadas. Elas são inseridas no Algoritmo Para-analisador para identificar o percentual da diferença entre as respostas dos dois Algoritmos (do programa em *Python* e do Para-analisador), obtendo assim maior assertividade na tomada de decisões.

Todos os níveis de dados pessoais não estruturados que podem transitar, isso somente tomando por base os *e-mails* de uma empresa impactam em vários níveis de uma organização com impacto diferenciado, por isso a preocupação crescente com esses dados torna-se cada vez mais presente diariamente, tanto a nível pessoal, quanto ao nível corporativo, pois o risco negativo pode ser desproporcional.

4.7 Definição das camadas internas

Conforme a Figura 14, primeira camada do DLP Paraconsistente é elaborada em *Python*. A segunda camada é oculta com a utilização do Algoritmo Para-Analisador. A terceira camada representa o resultado obtido.

Figura 14: Estrutura do DLP Paraconsistente:



Fonte: Autora

4.8 Construção da Base de Dados

Para qual cada uma das linhas capturadas apresentava uma descrição da origem dos dados disponibilizados pela empresa e anonimizados (Figura 15).

Figura 15: Amostra exemplo:

Nivel hierarquic	Dados	Fonte	Destino	Tamanho
3	DC	e-mail		1
3		e-mail	financeiro	1
3		e-mail	3	3
3		e-mail	administrativo	1
2	DC	e-mail	TI	1
2	DC	e-mail	administrativo	1
3	DC	e-mail	Imagem	996
1	DC	e-mail		1

Fonte: Autora

4.9 Avaliação da base de dados bruta pelo DLP - Paraconsistente

Em diversas empresas os agentes de tratamento (Controlador ou Processador) e o encarregado de dados (Data Protection Officer - DPO), ao analisar os dados não estruturados podem apresentar opiniões contraditórias, diferentes e inconsistentes, entretanto pode existir uma diferença sutil de entendimento quanto a validação desses dados, onde estes podem ser considerados válidos ou inconsistentes.

Neste caso, o universo da amostra utilizada no teste foi de 2000 dados não estruturados (e-mails), relativos a alertas cibernéticos obtidos por um *Security Information and Event Monitoring – SIEM* (Purview da Microsoft). O SIEM de mercado pré-seleciona os dados que considera alertas, mas gera muitos falsos positivos, e necessita que um analista especializado em LGPD avalie cada situação individualmente. Por isso, a proposta da primeira camada da ferramenta *Data Loss Prevention - DLP* Paraconsistente para validar essa massa de dados de forma bruta.

O programa desenvolvido em linguagem *Python* utilizou adaptações de algoritmos (KOPEC, 2019) e (MCKINNEY, 2022), parametrizado conforme necessidade do contexto:

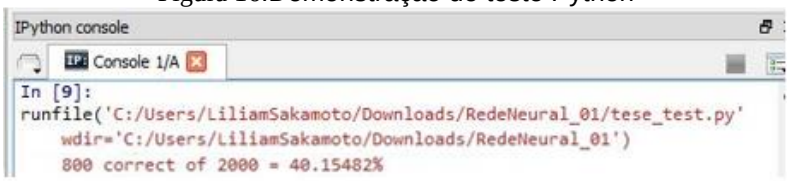
- Fonte *main.py*: Define o núcleo, para tratamento da leitura de arquivos da planilha com dados não estruturados;
- Fonte *util.py*: Define o tratamento de aprendizagem;
- Fonte *network.py*: Define o tratamento da saída pela Lógica Eτ;
- Fonte *layer.py*: Define o tratamento dos dados usando cálculos pela função de ativação, aprendizagem;
- Fonte *nft.py*: Define a efetuação de cálculos matemáticos com tratamento em matrizes, orientações pela função de ativação até obter resultados;
- Bibliotecas *Python* utilizadas: *import pandas as pad, import, import numpy as nup, import os, from typing, import List, from math.*

O arquivo anonimizado ao ser inserido e avaliado pelo programa em Python define quais os dados que apresentam os graus de incerteza e certeza associados a (μ, λ) definidos conforme ABE, 2015:

- Grau de incerteza: G_{un} (μ, λ) = μ + λ- 1 (0 ≤ μ, λ ≤ 1) e;
- Grau de certeza: G_{ce} (μ, λ) = μ - λ (0 ≤ μ, λ ≤ 1).

Observou-se que 40% desses dados foram considerados válidos (conforme Figura 5) e 60% inconsistentes, ou seja, uma grande quantidade de dados que podem ser descartados em situação normal, entretanto com o DLP Paraconsistente se tem a possibilidade de uma revalidação desses dados inconsistentes. Segue a demonstração do teste na Figura 16:

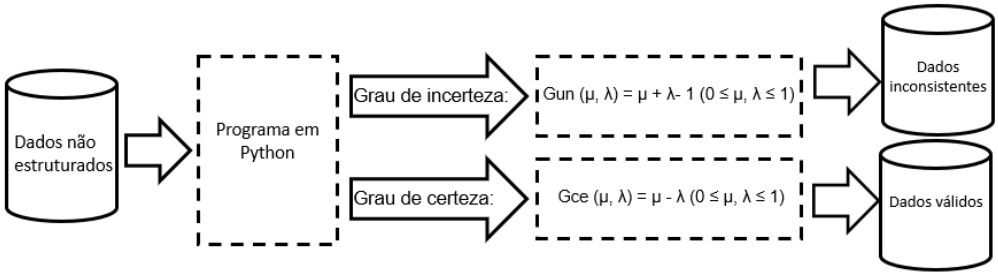
Figura 16:Demonstração do teste Python



Fonte: Autora

O algoritmo da primeira camada do DLP Paraconsistente pode ser demonstrado assim, conforme Figura 17:

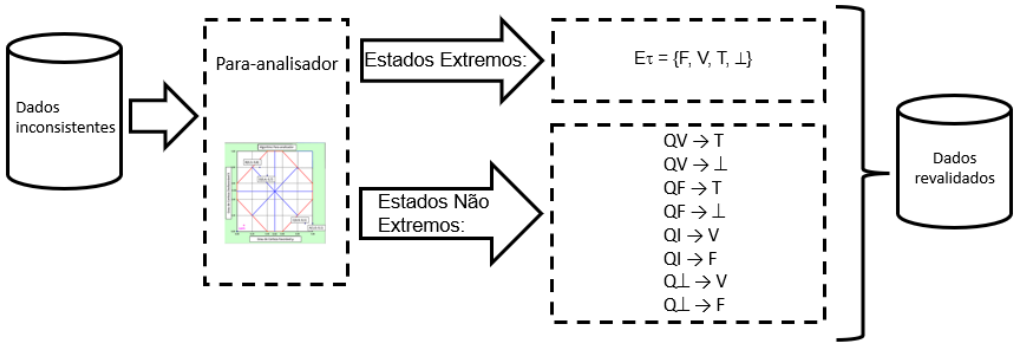
Figura 17: Algoritmo da primeira camada do DLP Paraconsistente



Fonte: Autora

Algoritmo da segunda camada do DLP Paraconsistente pode ser demonstrado assim, conforme Figura 18:

Figura 18: Algoritmo da segunda camada do DLP Paraconsistente



Fonte: Autora

4.10 Caracterização dos dados revalidados

Segue a seguir a caracterização dos estados extremos e não extremos para os dados revalidados após passarem pelo Para-analisador:

4.10.1 Estado Verdadeiro:

Caracterização do estado extremo Verdadeiro, conforme ABE et al (2011):

- G_{in} (μ, λ) = μ – λ = G_{ve} (μ, λ) ≥1/2 e μ ≥1/2 e λ≤1/2.

A tabela 10 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 2 – Seção 3 - Incidente de Privacidade de Dados é considerada verdadeira.

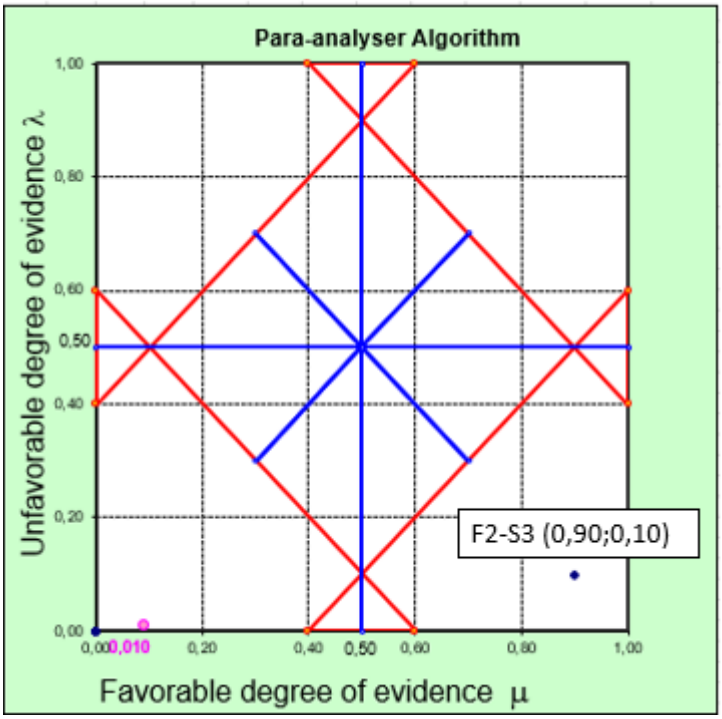
Tabela 10: Resultado do dado extremo revalidado Verdadeiro

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
D1	F2	S3	0,9	0,1	VERDADEIRO

Fonte: Autora

A Figura 19 demonstra o dado considerado inconsistente anteriormente e agora revalidado como Verdadeiro.

Figura 19: Demonstração de exemplo de dado revalidado como Verdadeiro



Fonte: Autora

4.10.2 Estado Falso:

Caracterização do estado Falso, conforme ABE et al (2011):

- $G_{in}(\mu, \lambda) = \mu - \Lambda = G_{fa}(\mu, \lambda) \leq 1/2$ e $\mu \leq 1/2$ e $\lambda \geq 1/2$

A tabela 11 a seguir demonstra exemplo dos resultados após passar pelo Para-analisador, onde o Fator 2 – Seção 1 - Suspeita de ataque de assinatura - tipo "*share*" são consideradas falsas.

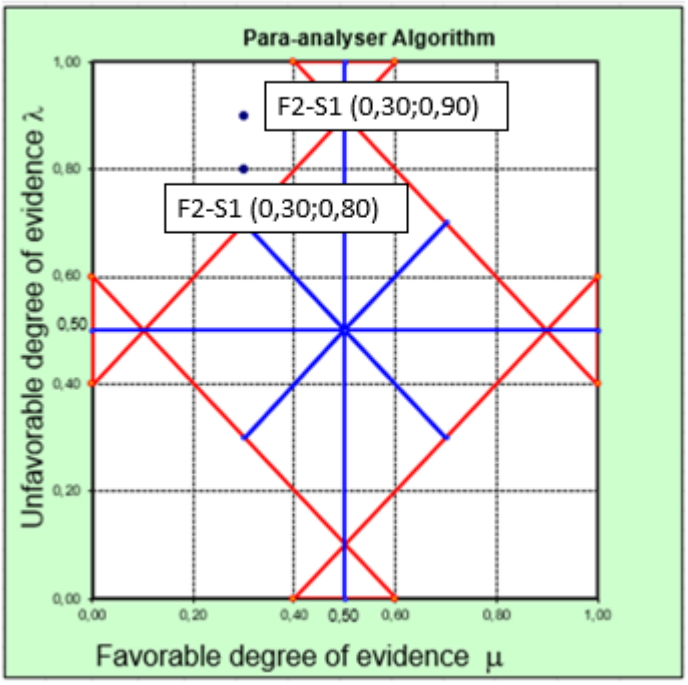
Tabela 11: Resultado do dado extremo revalidado Falso

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
B1	F2	S1	0,3	0,9	FALSO
B2	F2	S1	0,3	0,8	FALSO

Fonte: Autora

Estes exemplos estão demonstrados na Figura 20:

Figura 20: Estado Falso para Fator 2 – Seção 1:



Fonte: Autora

4.10.3 Estado Incompleto:

Caracterização do estado Incompleto, conforme ABE et al (2011):

- $G_{ce}(\mu, \lambda) = \mu + \lambda - 1 = G_{ct}(\mu, \lambda) \geq 1/2$ e $\mu \geq 1/2$ e $\lambda \geq 1/2$

A tabela 12 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 1 – Seção 1 - Comunicação com Rede TOR é considerada incompleta.

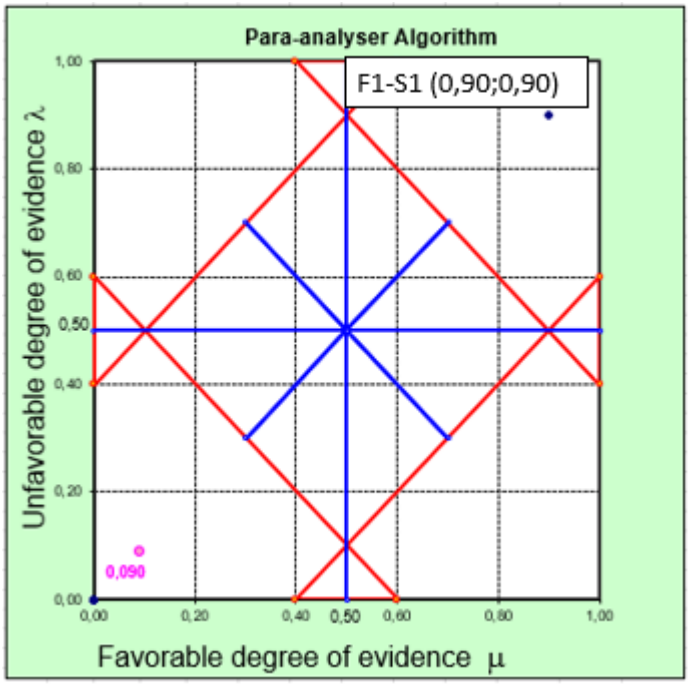
Tabela 12: Resultado para o Estado extremo revalidado Incompleto.

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
A1	F1	S1	0,9	0,9	INCOMPLETO

Fonte: Autora

Este exemplo está demonstrado na Figura 21:

Figura 21: Estado Incompleto para Fator 1 – Seção 1:



Fonte: Autora

4.10.4 Estado Paracompleto:

Caracterização do estado Paracompleto, conforme ABE et al (2011):

- $G_{ce}(\mu, \lambda) = \mu + \lambda - 1 = G_{pa}(\mu, \lambda) \leq -1/2$ e $\mu \leq 1/2$ e $\lambda \leq 1/2$

A tabela 13 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 1 – Seção 1 - Comunicação com Rede TOR é considerada incompleta.

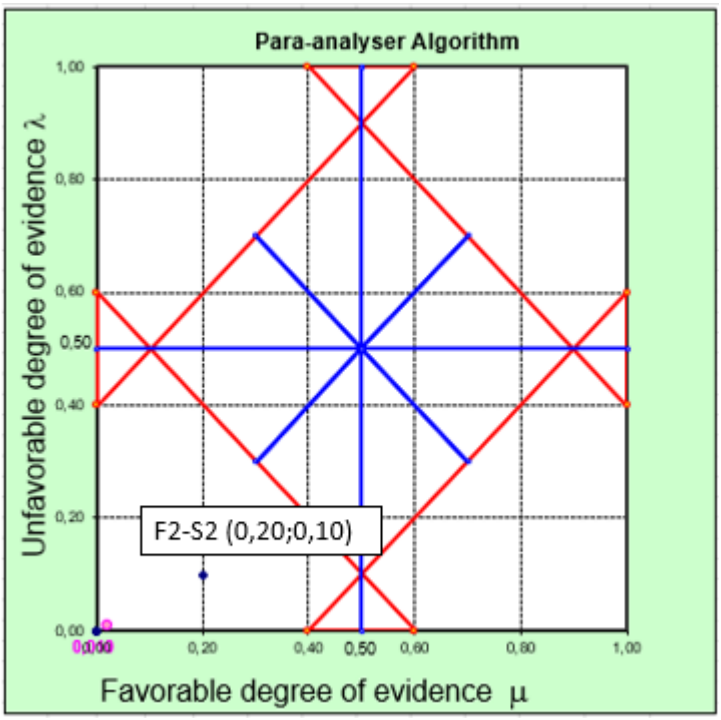
Tabela 13: Resultado do dado extremo revalidado Paracompleto

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
C1	F2	S2	0,2	0,1	PARACOMPLETO

Fonte: Autora

Este exemplo está demonstrado na Figura 22:

Figura 22: Estado Paracompleto para Fator 2 – Seção 2



4.10.5 Estado Quase Verdadeiro tendendo ao Inconsistente:

Caracterização do estado não extremo Quase Verdadeiro tendendo ao Inconsistente, conforme ABE et al (2011):

- $G_{in}(\mu, \lambda) = \mu - \lambda = G_{ve}(\mu, \lambda) \leq 1/2$ e $\mu \geq 1/2$ e
- $G_{ce}(\mu, \lambda) = \mu + \lambda - 1 = G_{ct}(\mu, \lambda) \geq 0$

A tabela 14 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 2 – Seção 3 - Suspeita de Incidente de Privacidade de Dados é considerada Quase Verdadeiro tendendo ao Inconsistente.

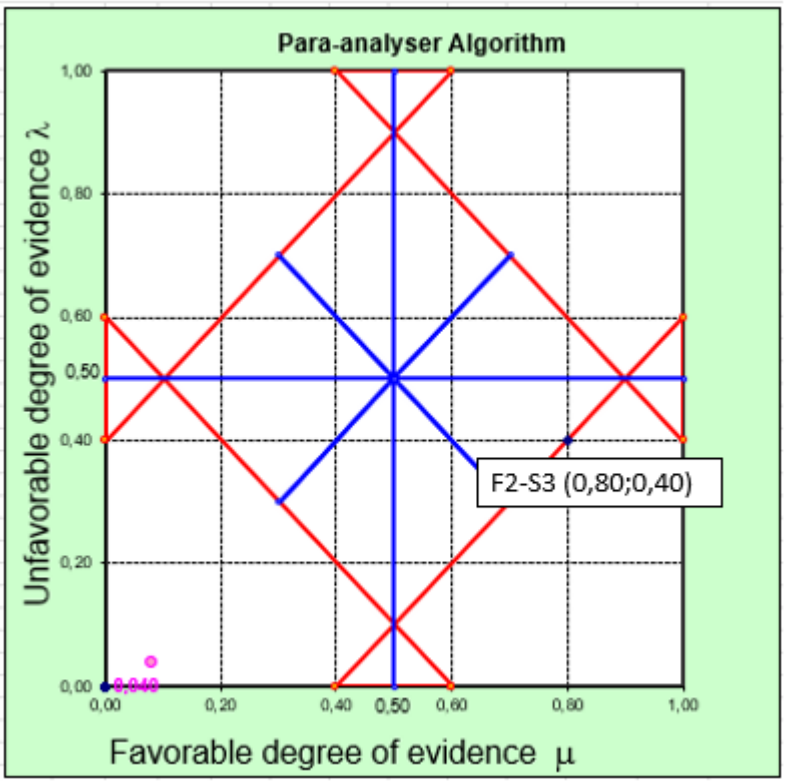
Tabela 14: Resultado do dado não extremo revalidado Quase Verdadeiro tendendo ao Inconsistente

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
D3	F2	S3	0,8	0,4	Quase Verdadeiro tendendo ao Inconsistente

Fonte: Autora

Este exemplo está demonstrado na Figura 23:

Figura 23: Estado não extremo Quase Verdadeiro tendendo ao Inconsistente para Fator 2 – Seção 3



Fonte: Autora

4.10.6 Estado Quase Verdadeiro tendendo ao Paracompleto:

Caracterização do estado não Quase Verdadeiro tendendo ao Paracompleto, conforme ABE et al (2011):

- G_{in} (μ, λ) = μ- λ = G_{ve} (μ, λ) ≤1/2 e μ ≥1/2 e
- G_{ce} (μ, λ) = μ+ λ-1 = G_{ct} (μ, λ) ≥0

A tabela 15 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 2 – Seção 3 - Suspeita de Incidente de Privacidade de Dados é considerada Quase Verdadeiro tendendo ao Paracompleto.

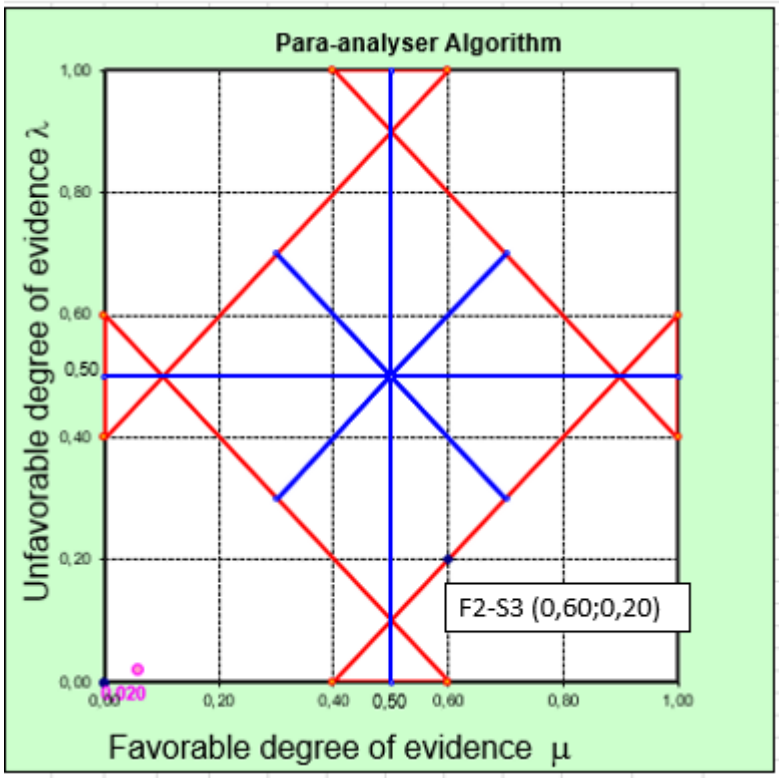
Tabela 15: Resultado do dado não extremo revalidado Quase Verdadeiro tendendo ao Paracompleto

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
D4	F2	S3	0,8	0,1	Quase Verdadeiro tendendo ao Paracompleto

Fonte: Autora

Este exemplo está demonstrado na Figura 24:

Figura 24: Estado não extremo Quase Verdadeiro tendendo ao Paracompleto para Fator 2 – Seção 3



Fonte: Autora

4.10.7 Estado Quase Falso tendendo ao Inconsistente:

Caracterização do estado não extremo Quase Falso tendendo ao Inconsistente, conforme ABE et al (2011):

- $G_{in}(\mu, \lambda) = \mu - \lambda = G_{fa}(\mu, \lambda) \geq -1/2$ e $\mu \leq 1/2$ e
- $G_{ce}(\mu, \lambda) = \mu + \lambda - 1 = G_{ct}(\mu, \lambda) \geq 0$

A tabela 16 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 2 – Seção 1 - Suspeita de ataque de assinatura - tipo "*share*" é considerada Quase Falso tendendo ao Inconsistente.

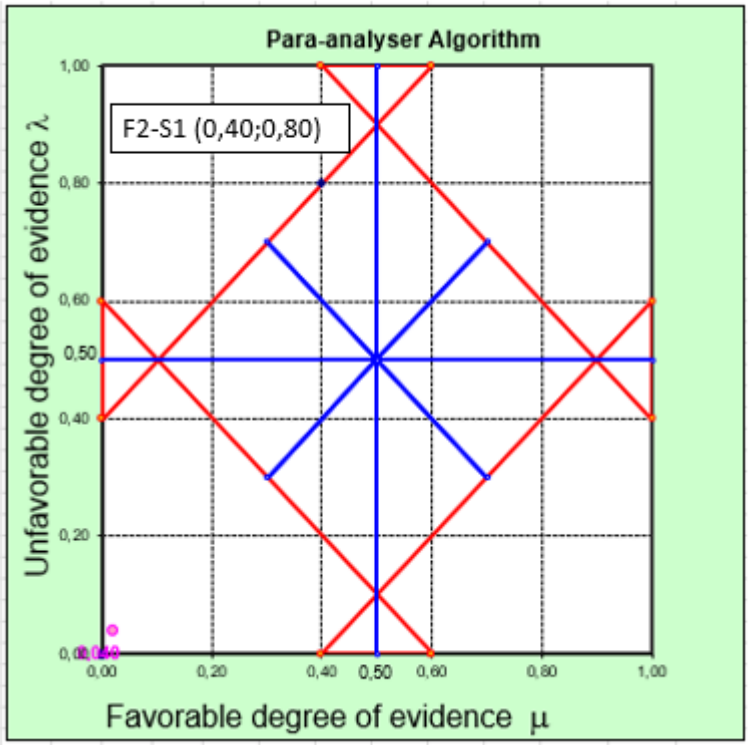
Tabela 16: Resultado do dado não extremo revalidado Quase Falso tendendo ao Inconsistente

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
B3	F2	S1	0,4	0,8	Quase Falso tendendo ao Inconsistente

Fonte:Autora

Este exemplo está demonstrado na Figura 25:

Figura 25: Estado não extremo Quase Falso tendendo ao Inconsistente para Fator 2 – Seção 1



Fonte: Autora

4.10.8 Estado Quase Falso tendendo ao Paracompleto:

Caracterização do estado não extremo Quase Falso tendendo ao Paracompleto, conforme ABE et al (2011):

- G_{in} (μ, λ) = μ- λ = G_{fa} (μ, λ) ≥-1/2 e μ ≤1/2 e
- G_{ce} (μ, λ) = μ+ λ-1 = G_{ct} (μ, λ) ≥0

A tabela 17 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 1 – Seção 1 - Comunicação com Rede TOR é considerada Quase Falso tendendo ao Paracompleto.

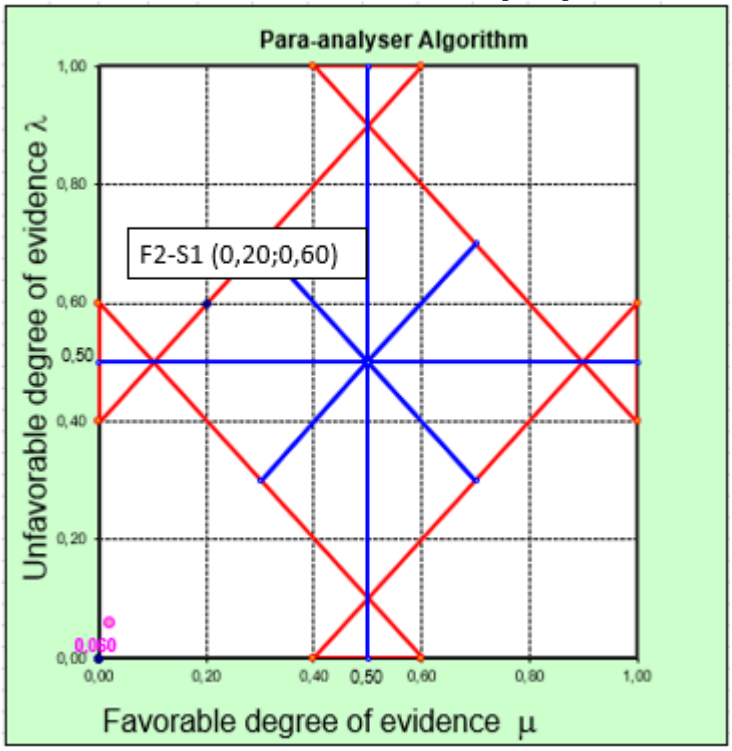
Tabela 17: Resultado do dado não extremo revalidado Quase Falso tendendo ao Paracompleto

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
A4	F1	S1	0,2	0,6	Quase Falso tendendo ao Paracompleto

Fonte: Autora

Este exemplo está demonstrado na Figura 26:

Figura 26: Estado não extremo Quase Falso tendendo ao Paracompleto para Fator 2 – Seção 1



Fonte: Autora

4.10.9 Estado Quase Inconsistente tendendo ao Verdadeiro:

Caracterização do estado não extremo Quase Inconsistente tendendo ao Verdadeiro, conforme ABE et al (2011):

- $G_{ce}(\mu, \lambda) = \mu + \lambda - 1 = G_{ct}(\mu, \lambda) \leq 1/2$ e $\lambda \geq 1/2$ e
- $G_{in}(\mu, \lambda) = \mu - \lambda = G_{ve}(\mu, \lambda) \geq 0$

A tabela 18 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 1 – Seção 1 - Comunicação com Rede TOR é considerada Quase Falso tendendo ao Paracompleto.

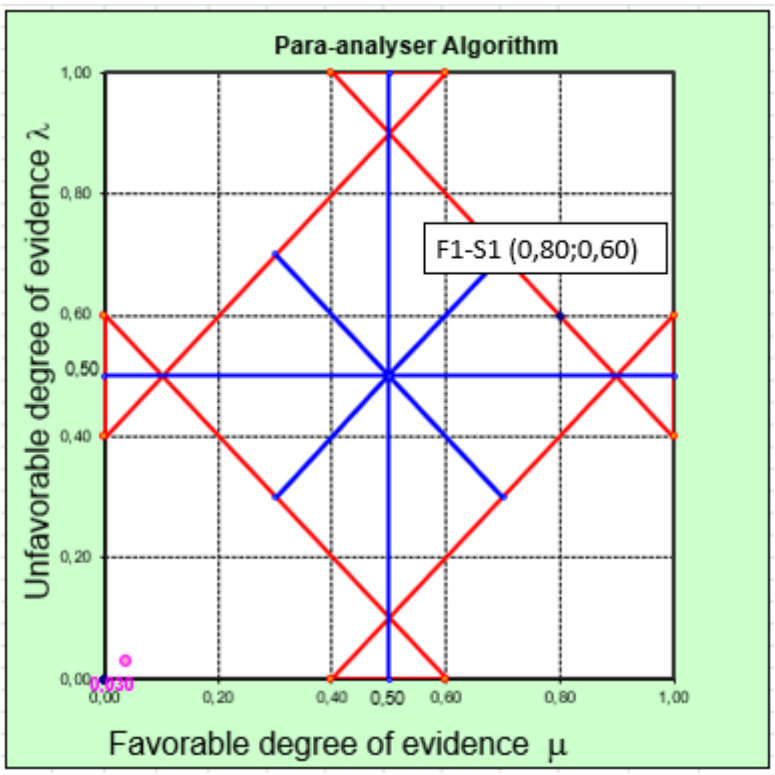
Tabela 18: Resultado do dado não extremo revalidado Quase Inconsistente tendendo ao Verdadeiro

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
A5	F1	S1	0,8	0,6	Quase Inconsistente tendendo ao Verdadeiro

Fonte: Autora

Este exemplo está demonstrado na Figura 27:

Figura 27: Estado não extremo Quase Inconsistente tendendo ao Verdadeiro para Fator 1 – Seção 1



Fonte: Autora

4.10.10 Estado Quase Inconsistente tendendo ao Falso:

Caracterização do estado não extremo Quase Inconsistente tendendo ao Falso, conforme ABE et al (2011):

- $G_{ce}(\mu, \lambda) = \mu + \lambda - 1 = G_{ct}(\mu, \lambda) \leq 1/2$ e $\mu \geq 1/2$ e
- $G_{in}(\mu, \lambda) = \mu - \lambda = G_{fa}(\mu, \lambda) \leq 0$

A tabela 19 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 1 – Seção 1 - Comunicação com Rede TOR é considerada Quase Inconsistente tendendo ao Falso.

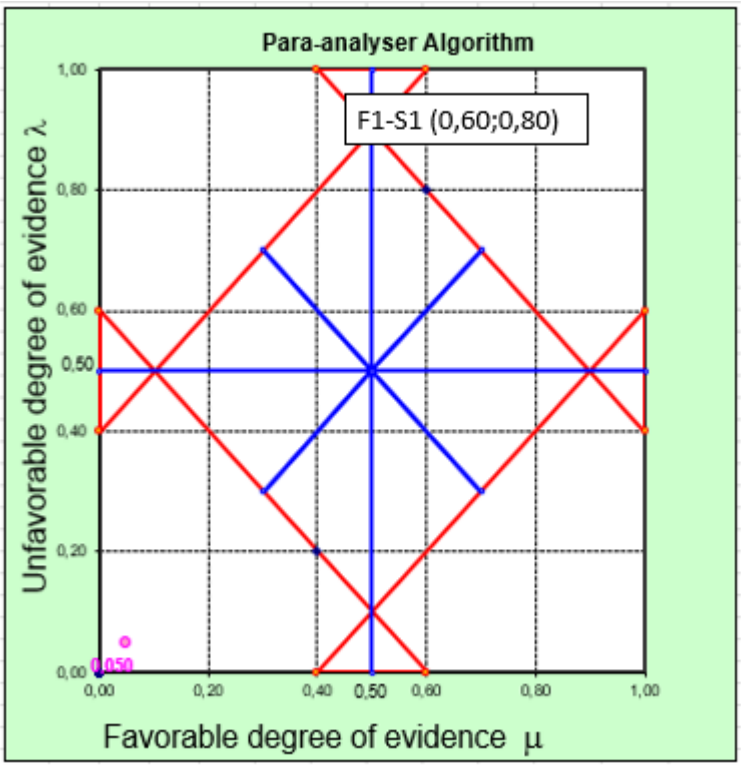
Tabela 19: Resultado do dado não extremo revalidado Quase Inconsistente tendendo ao Falso

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
A6	F1	S1	0,6	0,8	Quase Inconsistente tendendo ao Falso

Fonte: Autora

Este exemplo está demonstrado na Figura 28:

Figura 28: Estado não extremo Quase Inconsistente tendendo ao Falso para Fator 1 – Seção 1



Fonte: Autora

4.10.11 Estado Quase Paracompleto tendendo ao Verdadeiro:

Caracterização do estado não extremo Quase Paracompleto tendendo ao Verdadeiro, conforme ABE et al (2011):

- $G_{ce}(\mu, \lambda) = \mu + \lambda - 1 = G_{pa}(\mu, \lambda) \geq -1/2$ e $\lambda \leq 1/2$ e
- $G_{in}(\mu, \lambda) = \mu - \lambda = G_{ve}(\mu, \lambda) \geq 0$

A tabela 20 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 2 – Seção 2 - Suspeita de ransomware – tipo “*create*” é considerada Quase Paracompleto tendendo ao Verdadeiro.

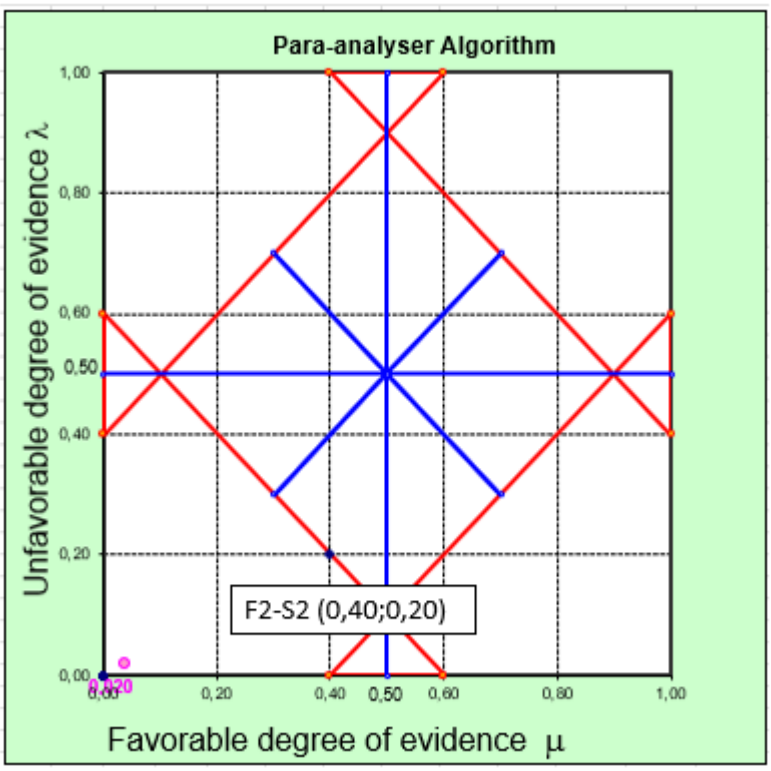
Tabela 20: Resultado do dado não extremo revalidado Quase Paracompleto tendendo ao Verdadeiro

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
C2	F2	S2	0,4	0,2	Quase Paracompleto tendendo ao Verdadeiro

Fonte: Autora

Este exemplo está demonstrado na Figura 29:

Figura 29: Estado não extremo Quase Paracompleto tendendo ao Verdadeiro para Fator 2 – Seção 2



Fonte: Autora

4.10.12 Estado Quase Paracompleto tendendo ao Falso:

Caracterização do estado não extremo Quase Paracompleto tendendo ao Falso, conforme ABE et al (2011):

- $G_{in}(\mu, \lambda) = \mu + \lambda - 1 = G_{pa}(\mu, \lambda) \geq -1/2$ e $\lambda \leq 1/2$ e

- $G_{in}(\mu, \lambda) = \mu - \lambda = G_{fa}(\mu, \lambda) \leq 0$

A tabela 21 a seguir demonstra o resultado após passar pelo Para-analisador, onde o Fator 2 – Seção 2 - Suspeita de ransomware – tipo “*create*” é considerada Quase Paracompleto tendendo ao Falso.

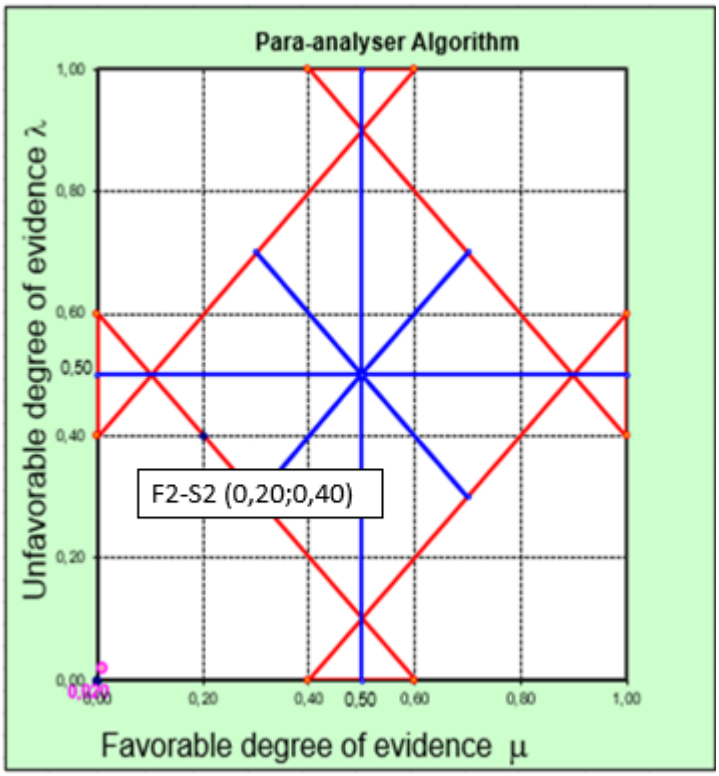
Tabela 21: Resultado do dado não extremo revalidado Quase Paracompleto tendendo ao Falso

Resultado	Fator	Seção	Conclusão		Definição
			μ	λ	
C3	F2	S2	0,2	0,4	Quase Paracompleto tendendo ao Falso

Fonte: Autora

Este exemplo está demonstrado na Figura 30:

Figura 30: Estado não extremo Quase Paracompleto tendendo ao Falso para Fator 2 – Seção 2



Fonte: Autora

4.11 Definição dos aspectos de tomada de decisão

Entretanto, o intervalo considerado falso pode passar um crivo maior que necessitariam de um monitoramento mais efetiva, devido ao percentual de ocorrências que poderia ser um alerta de vazamento de dados de forma sutil.

A massa de dados capturada como amostragem inicial, apresentava cerca de 30 informações por dia durante 1 mês, pois por dia são gerados cerca de 1.000 a 2.000 linhas para análise para o encarregado de dados. A amostra compôs um montante para análise de 2.000 dados, dos quais foram verificadas pelo Algoritmo em *Python* como efetiva o percentual de 60% como contraditório.

Os dados considerados contraditórias foram repassados no Algoritmo Para-analisador e obteve-se um diferencial de 9% verdadeiro, 6% incompleto, 10% paracompleto e 15% como falso. Totalizando 24% de otimização dos dados que seriam descartados (9% verdadeiro e 15% falso).

CAPÍTULO 5 – RESULTADOS OBTIDOS

5.1 Resultados apresentados nesta tese

A massa de dados da amostra capturada foi de 2.000 dados, dos quais foram verificadas pelo Algoritmo em *Python* como efetiva o percentual de 60% como contraditório.

Os dados considerados contraditórias foram repassados no DLP - Paraconsistente e obteve-se um diferencial de 9% verdadeiro, 6% incompleto, 10% paracompleto e 15% como falso. Totalizando 24% de otimização dos dados que seriam descartados (9% verdadeiro e 15% falso).

Conclui-se que os estudos do DLP paraconsistente para prevenção de perda/dano/roubo/evasão de dados pessoais não estruturados com foco na Lei LGPD otimiza o monitoramento de incidentes e assertividade com a legislação, pois 60% dos dados seriam descartados, enquanto com o uso desse estudo foram aproveitados 24%, diminuindo o descarte somente para 36%.

Situação que otimiza as atividades do encarregado de dados em sua performance para proteção e privacidade de dados não estruturados.

5.2 Resultados apresentados nos artigos

São apresentados nos artigos elaborados durante o curso com os detalhes sobre publicações e submissões:

Artigo publicado referente ao objetivo geral:

- O objetivo geral desta pesquisa é estudar a otimização da detecção de perda de dados através de *Data Loss Prevention* – DLP Paraconsistente para prevenção de dano/roubo/vazamento de dados pessoais não estruturados com foco na Lei Geral de Proteção de Dados - LGPD para auxílio aos encarregados de dados ou *Data Protection Officer - DPO*.
    - Artigo A: *Professional Guidance of the DPOs – BR in Corporate Governance in Logistics Chains*, neste artigo aborda a importância

de um guia de treinamento para os DPOs (Apêndice I), apresentado no Congresso APMS 2022 e publicado em: https://link.springer.com/chapter/10.1007/978-3-031-16411-8_8.

Artigos publicado e submetidos referentes aos objetivos específicos:

- Prover identificação de perda/dano/roubo/vazamento de dados ou *Data Loss Prevention – DLP* Paraconsistente:
    - Identificar perda de dados pessoais não estruturados (clientes, funcionários e fornecedores - incidentes LGPD):
        - Artigo B: apresentado e publicado no 9º Congresso da Turquia *ÇUKUROVA – Overview of DPOs in the use of LGPD compliance software with Logic Paraconsistent Annotated Evidential $E\tau$ and DLP – Data Loss Prevention*, onde foi aplicado no estudo de software para LGPD (Apêndice II), publicado o *book* de *full text* em: https://en.iksadkongre.net/kongre-kitaplari.
    - Identificar perda de dados corporativos não estruturados (rede, armazenamento, *endpoint* e *cloud*):
        - Artigo C: *Optimizing the Data Loss level using Logic Paraconsistent Annotated Evidential $E\tau$* (Apêndice III), submetido para revista A1 – *World Development.*
- Prover o detalhamento da estrutura do *Data Loss Prevention* - DLP Paraconsistente;
    - Análise de segurança de dados no ambiente:
        - Artigo D: *Metaverse security using DLP and Paraconsistent Logic* (Apêndice IV), submetido para revista A1 - *Journal of Management in Engineering*.
        - Artigo E: *DLP: prevención de pérdida de datos con lógica paraconsistente para la seguridad en el metaverso*

(Apêndice VI), submetido para revista A1 - *Enseñanza de las Ciencias*.

- Aplicar *Data Loss Prevention* - DLP Paraconsistente para análise de perda/dano/roubo/vazamento de dados não estruturados.
    - Artigo F: Melhoria da utilização de um Sistema de Monitoramento de Eventos de Incidentes Cibernéticos com uso do DLP - Paraconsistente - *Optimization of SIEM using DLP and PANN* (Apêndice V), submetido para revista A1 – *IEEE System Journal.*

5.3 Detalhe dos resultados obtidos nos artigos

Artigo A: *Professional Guidance of the DPOs – BR in Corporate Governance in Logistics Chains*

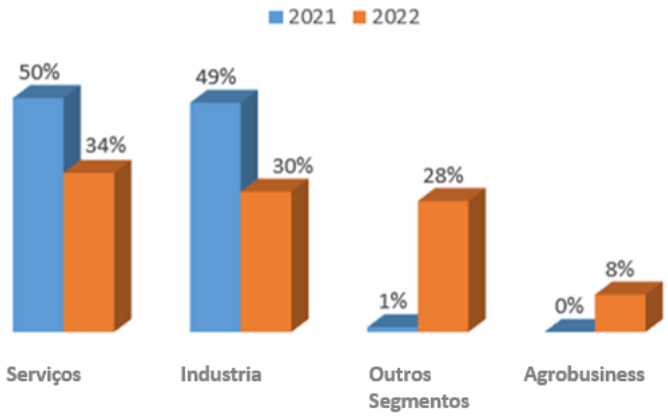
Foram analisados neste artigo uma amostra de 4.000 Encarregados de Dados ou *Data Protection Officer – DPO,* e verificados os dados para necessidade treinamento e capacitação profissional, divididos nos seguintes ramos de negócio em 2021:

- serviços (50%);
- comércio (17%);
- indústria (15%);
- outros segmentos (14%); e
- agronegócio (4%).

Artigo B: *Overview of DPOs in the use of LGPD compliance software with Logic Paraconsistent Annotated Evidential* $E\tau$ *and DLP – Data Loss Prevention*

Foram analisados neste artigo uma amostra de 8.000 Encarregados de Dados ou *Data Protection Officer* – DPO, sobre a atuação nas áreas de negócio, foram comparados os dados de 2021 e 2022, conforme gráfico da Figura 31:

Figura 31: Resultado do Artigo do 9º Congresso da Turquia *ÇUKUROVA*



Fonte: APDADOS 2021 e 2022

O encarregado de dados precisa utilizar uma combinação de diversas técnicas e ferramentas de análise para auxiliá-lo como DLP – *Data Loss Prevention*, aqui a análise é dentro das fases de desenvolvimento de software para compliance com a LGPD, que é segregada em quatro fases (Fase 1 – Dados Operacionais intuitivos, Fase 2 – Dados técnicos de Controles Internos Definidos, Fase 3 – Diferenciações Mensuráveis de Monitoramento e Fase 4 – Tomada de Decisão), entretanto neste artigo somente analisou-se os dados referente a fase 2.

Foram coletadas também na parte de pesquisa aplicada, 32 respostas de especialistas *Data Protection Officers - DPOs* para dois fatores analisados, para:

- Fator F1 – Há motivação nas empresas para adoção da Lei LGPD? No primeiro caso o resultado foi “verdadeiro” para a motivação da adoção do *software* para LGPD, validado ao se comparado a estatística do ano 2022 (amostra de 4.000 Encarregados de Dados);
- Fator F2 - Existem práticas formalizadas de Proteção de Dados e Privacidade? No segundo caso, o resultado o resultado foi inconsistente, pois o *software* não apresentava compliance com consistência na formalização prática de Proteção e Privacidade de dados, sendo que não existiam dados comparativos estatísticos para análise desse item.

Artigo C: *Optimizing the Data Loss level using Logic Paraconsistent Annotated Evidential Eτ*

A massa de dados utilizada anonimizada foi de empresa financeira que apresentava cerca de 30 informações por dia durante um mês, totalizando um montante para análise de 1.107 dados, dos quais o algoritmo verificou 60% em Python como eficazes e 40% como contraditórios.

Os dados considerados contraditórios foram repassados ao algoritmo do Para-analisador, obtendo-se um diferencial de 9% verdadeiros, 6% incompletos, 10% paracompletos e 15% falsos. Ou seja, foram aproveitados 24% dos dados que seriam descartados (os falsos e os verdadeiros).

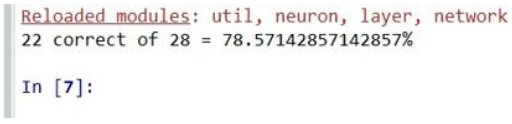
Porém, o intervalo considerado falso pode passar por um crivo mais excelente que exige um monitoramento mais eficaz devido ao percentual de ocorrências que poderiam ser um alerta de vazamento de dados de forma sutil.

Artigo D: *Metaverse security using DLP and Paraconsistent Logic*

Foi realizada pesquisa exploratória com a coleta de uma amostra anonimizada de 28 ativos referente a um mês de dados de uma empresa de transportes, e esses dados foram analisados por um programa em Python em conjunto com a Lógica Paraconsistente.

Realizado o teste em *Python* 78% foram identificados como correto e 22% desconsiderado por ser inconsistente, após pelo DLP Paraconsistente ocorreu uma otimização de 11%, conforme teste, vide figura 32:

Figura 32: Teste para o artigo *Metaverse security using DLP and Paraconsistent Logic.*



Fonte: autora.

Por meio do monitoramento, 22% foram desconsiderados neste processo, porém com a utilização do DLP - Paraconsistente, ocorreu uma otimização de 11% na análise de segurança.

Artigo E: *DLP: prevención de pérdida de datos con lógica paraconsistente para la seguridad en el metaverso*

Com DLP e banco de dados fornecido pela transportadora com 200 itens analisados. Verificou-se que uma quantidade significativa de dados seria descartada na primeira etapa do processo (37%) por não apresentarem uma definição operacional quanto à situação desses ativos. Com a utilização do DLP - Paraconsistente, observou-se que em comparação com a otimização para 23%, ou seja, utilização de mais de 15% dos dados.

Artigo F: *Optimization of SIEM using DLP and PANN*

Apresenta pesquisa exploratória com testes em um banco de dados anonimizado de 200 alertas de segurança cibernética de uma empresa de transportes durante 14 meses extraído de um SIEM focado em Antivírus. Nesta etapa, o estudo sugere a implementação de um DLP – Paraconsistente. Os resultados mostram que na análise pura do SIEM foram descartados 37% dos alertas considerados dados inconclusivos; porém, com a aplicação da solução sugerida, é possível atingir a minimização do descarte de apenas 20% dos alertas de cibersegurança, representando uma otimização de 17%.

CAPÍTULO 6 – CONSIDERAÇÕES FINAIS

6.1 Conclusões

O objetivo geral desta pesquisa foi atingido, pois se realizou de forma detalhada o estudo da otimização da detecção de perda de dados por meio do *Data Loss Prevention – DLP* Paraconsistente, conforme resultados obtidos:

- Na tese:
    - Conclui-se que os estudos do DLP paraconsistente para prevenção de perda/dano/roubo/evasão de dados alertas de segurança não estruturados com o uso desse estudo foram reaproveitados 24%, diminuindo o descarte inicial de 60% somente para 36%.
- E no artigo A: *Professional Guidance of the DPOs – BR in Corporate Governance in Logistics Chains*
    - Conclui-se da amostra analisada que a maior parte dos profissionais Encarregados de Dados, que estão alinhados com compliance e treinamento são da área de serviços (50%). Observando-se que o foco em pessoas é muito importante para o desenvolvimento dos Encarregados de Dados estarem preparados para minimizar a perda/vazamento/roubo de dados.

Os objetivos específicos foram atingidos, conforme apresentado nos resultados dos artigos:

- No artigo B: *Overview of DPOs in the use of LGPD compliance software with Logic Paraconsistent Annotated Evidential Eτ and DLP – Data Loss Prevention*
    - Conclui-se a importância do foco em *Compliance* com a LGPD:

        - Feita a comparação entre a amostra de dois anos quanto aos profissionais Encarregados de Dados quanto ao ramo de atuação, a área de serviços é o que mais se destaca: 50% em 2021 e 34% em 2022.
        - Quanto a delimitação inicial do DLP para avaliação de software de compliance com a LGPD:
            - Fator F1 – Há motivação nas empresas para adoção da Lei LGPD? Foi validado como “verdadeiro”, este fator tanto pelo *Data Loss Prevention - DLP* Paraconsistente, quanto pelas estatísticas do ano de 2022;
            - Fator F2 - Existem práticas formalizadas de Proteção de Dados e Privacidade? Não se chegou a uma resposta consistente, pois o software não apresentava compliance com a formalização prática de Proteção e Privacidade de dados.
- No artigo C: *Optimizing the Data Loss level using Logic Paraconsistent Annotated Evidential Eτ*
    - Conclui-se a aderência do foco na ferramenta com massa analisada pelo programa desenvolvido na linguagem em Python que identifica os dados válidos e os contraditórios de acordo com o Grau de Evidência Favorável e o Grau de Evidência Desfavorável respectivamente, sendo que dos 40% dos dados considerados contraditórios inicialmente, foram reaproveitados 24% dos dados que seriam descartados.
- No artigo D: *Metaverse security using DLP and Paraconsistent Logic*
    - Conclui-se o foco do detalhamento da infraestrutura e dos testes, que por meio do monitoramento simples ou uso somente do DLP (Programa em Python), onde 22% foram desconsiderados neste processo, porém com a utilização do

DLP - Paraconsistente, a minimização foi de 22%, ou seja, uma otimização de 11% na análise de segurança.

- No artigo E: *DLP: prevención de pérdida de datos con lógica paraconsistente para la seguridad en el metaverso*
    - Conclui-se que se validou o foco nos testes com uma quantidade maior de massa de dados utilizando-se o DLP - Paraconsistente, observou-se que em comparação com uma otimização para 23%, ou seja, um reaproveitamento de 15% dos dados a mais.
- No artigo F: *Optimization of SIEM using DLP and PANN*
    - Conclui-se o foco na aplicação prática do DLP Paraconsistente em cibersegurança, onde os resultados anteriormente apresentados pelo SIEM de mercado foram descartados 37% dos alertas considerados dados inconclusivos; porém, com a aplicação do DLP - Paraconsistente, é possível atingir a minimização do descarte de apenas 20% dos alertas de cibersegurança, representando uma otimização de 17%.

Discussão:

Neste estudo, o foco foi somente quanto aos dados não estruturados, onde especificamente testamos uma base de dados de e-mails.

Analisando que estudos futuros necessitam de maior aprofundamento com relação principalmente aos tipos de dados corporativos e que são estruturados. Pode-se abordar os tipos de armazenamento dos dados estruturados, como por exemplo, quais estão na nuvem e quais estão nos *endpoints*, ou seja, com os usuários finais em *desktops*, *notebooks*, *tablets*, smartfones, *smartwatches* etc. Como este foi o foco da discussão pode-se observar que a LGPD também aborda a necessidade de verificar os dados que estão armazenados em meios físicos, como relatórios impressos, microfilmes e arquivos mortos não digitalizados. Entretanto, não haveria tempo hábil para um estudo aprofundado devido ao volume de dados para análise ser muito extenso e variado.

As áreas para estudos futuros são descritas na figura 33:

Figura 33: Áreas discutidas para análise de dados estruturados

Fonte: Autora

6.2 Sugestão de trabalhos futuros:

A previsão de trabalhos futuros está na aplicação prática de DLP paraconsistente para dados estruturados e não estruturados também em vários ramos que necessitam desse alinhamento, como:

- Farmácias,
- Saúde,
- Contabilidade,
- Financeiro,
- Educação e
- Sustentabilidade.

Existindo a necessidade de adequação e melhoria do modelo para o direcionamento da tomada de decisão.

Além da utilização da visão das bases legais abordadas pela LGPD para cada uma das áreas acima citadas:

- Consentimento do titular;
- Obrigações legais;
- Execuções de políticas públicas;

- Estudos de órgão de pesquisa;
- Execução de contratos;
- Exercício de regular diretos em investigação judicial;
- Proteção de vida;
- Tutela da saúde;
- Interesses legítimos do controlador ou de terceiros;
- Proteção ao crédito.

REFERÊNCIAS

ABE, Jair Minoro. (1992) Fundamentos da lógica anotada. Tese de Doutorado. Universidade de São Paulo. 1992

ABE, J.M., et al. (2011) Lógica Paraconsistente Anotada Evidencial Et, pp. 38–39. Comunnicar, Santos, 2011.

ABE, Jair Minoro (Ed.). (2015) Paraconsistent intelligent-based systems: New trends in the applications of paraconsistency (Vol. 94). Springer, 2015.

ABE, J. M., AKAMA, S., & NAKAMATSU, K. (2015). Introduction to annotated logics: foundations for paracomplete and paraconsistent reasoning (Vol. 88). Springer. AHMAD, Shafi et al. Microsoft Purview: A System for Central Governance of Data. Proceedings of the VLDB Endowment, v. 16, n. 12, p. 3624-3635, 2023.

AKAMA, S. (Ed.). (2016). Towards Paraconsistent Engineering. Cham: Springer International Publishing.

ALABADI, Montdher; CELIK, Yuksel. (2020) Anomaly detection for cyber-security based on convolution neural network: A survey. In: 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA). IEEE, 2020. p. 1-14.

ANDRADE, Maria Margarida de. (2018) Introdução à metodologia do trabalho científico: elaboração de trabalhos na graduação. 10. ed. São Paulo: Atlas,2018.

ANPD – Autoridade Nacional de Proteção de Dados (2023) Regulamento de Dosimetria Disponível em: https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-publica-regulamento-de-dosimetria. Acessado em 20/05/2023.

BACH-NUTMAN, Matthew. (2020) Understanding the top 10 owasp vulnerabilities. arXiv preprint arXiv:2012.09960, 2020.

BRASIL. Constituição Federal, de 1988. (1988) Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acessado em: 09/02/2023.

BRASIL. Lei Geral de Proteção de Dados Pessoais (LGPD). Lei nº 13.709, de 14 de agosto de 2018. (2018) Disponível em:http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm . Acessado em: 21/04/2023.

BRASIL, 1996 – Lei da Arbitragem – LEI Nº 9.307 de 23 de setembro de 1996. (1996) Disponível em: https://legislacao.presidencia.gov.br/atos/?tipo=LEI&numero=9307&ano=1996&ato=121IzZq1UMJpWT25d. Acessado em: 27/11/2023.

CALVIN, Christopher; EULERICH, Marc; HOLT, Matthew. (2023) Characteristics of Cybersecurity and it Involvement by the Ia Activity. Available at SSRN 4572560, 2023.

CBO - Classificação Brasileira de Ocupações – 1421-35 (mtecbo.gov.br). (2023) Disponível em: https: https://www.mtecbo.gov.br/cbosite/pages/pesquisas/BuscaPorTituloResultado.jsf. Acessado em: 05 fev. 2023.

CINQUE, Marcello; COTRONEO, Domenico; PECCHIA, Antonio.(2018) Challenges and directions in security information and event management (SIEM). In: 2018 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW). IEEE, 2018. p. 95-99.

COOPER, Donald R.; SCHINDLER, Pamela S. (2016) Métodos de pesquisa em administração. 12. ed. Porto Alegre: AMGH, 2016.

DA SILVA FILHO, João Inácio. (1999) Métodos de Aplicações da Lógica Paraconsistente Anotada de Anotação com dois Valores -LPA2v com Construção de Algoritmo ET. Universidade de São Paulo, 1999.

DA SILVA FILHO, João Inácio. (2008) Lógica paraconsistente e probabilidade pragmática no tratamento de incertezas. Seleção Documental: Inteligência Artificial e novas Tecnologias, n. 9, p. 16-27, 2008.

DA SILVA FILHO, João Inácio; TORRES, Germano Lambert; ABE, Jair Minoro. Uncertainty treatment using paraconsistent logic: Introducing paraconsistent artificial neurais networks. IOS Press, 2010.

DE CARVALHO, Fábio Romeu; ABE, Jair Minoro. (2018) A Paraconsistent Decision-Making Method. Springer International Publishing, 2018.

DONG, Shi; ABBAS, Khushnood; JAIN, Raj. (2019) A survey on distributed denial of service (DDoS) attacks in SDN and cloud computing environments. IEEE Access, v. 7, p. 80813-80828, 2019.

EBERENDU, Adanma Cecilia et al. (2016) Unstructured Data: an overview of the data of Big Data. International Journal of Computer Trends and Technology, v. 38, n. 1, p. 46-50, 2016.

EUROPEAN COMISSION - GDPR - General Data Protection Regulation 2016 (2016) Disponível em: < https://gdpr-info.eu/ > Acessado em: 11/02/2023.

EUROPEAN COMISSION - Guidelines on Data Protection Officers ('DPOs') (wp243rev.01) (2016) Disponível em:<https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612048> Acessado em: 21/04/2023.

EUROPEAN COMISSION Guidelines on Consent under Regulation 2016/679 (wp259rev.01) (2016) Disponível em:<https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051 > Acessado em: 20/05/2023.

EUROPEAN COMISSION GDPR - General Data Protection Regulation (2016) Disponível em:< https://gdpr-info.eu/ > Acessado em: 19/03/2023.

KHAN, Freeha et al. (2021) Data breach management: an integrated risk model. Information & Management, v. 58, n. 1, p. 103392, 2021.

KOPEC, D. (2019). Classic computer science problems in Python. Simon and Schuster.

GIL, Antônio Carlos. (2018) Como elaborar projetos de pesquisa. 6. ed. São Paulo: Atlas, 2018.

HART, Michael; MANADHATA, Pratyusa; JOHNSON, Rob. (2011) Text classification for data loss prevention. In: International Symposium on Privacy Enhancing Technologies Symposium. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. p. 18-37.

HELLEBRAND, Hans-Martin. (2017) An Exponential World: Nature, Patterns, and How to Leverage Them. The Palgrave Handbook of Managing Continuous Business Transformation, p. 95-113, 2017.

HINDS, Joanne; WILLIAMS, Emma J.; JOINSON, Adam N. (2020) “It wouldn't happen to me”: Privacy concerns and perspectives following the Cambridge Analytica scandal. International Journal of Human-Computer Studies, v. 143, p. 102498, 2020.

HORODYSKI, Dominik. (2014) 2013 OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data as an Example of Recent Trends in Personal Data Protection. Collective human rights, p. 255, 2014.

JANSEN, Rob et al. (2012)Methodically Modeling the Tor Network. In: CSET. 2012.

LIU, Tzu-Hsin; HUNG, Shih-Chang; CHU, Yee-Yeen. (2007) Environmental jolts, entrepreneurial actions and value creation: A case study of Trend Micro. Technological forecasting and social change, v. 74, n. 8, p. 1432-1445, 2007.

LIMA, Luiz A. Comitê Científico APDADOS Estatísticas: Panorama da Conscientização Nacional sobre LGPD. CNPPD 2022. Disponível em: https://cnppd.online/. Acessado em: 24/08/2022.

MALDONADO, Viviane Nóbrega; BLUM, Renato Ópice. (2022) LGPD: Lei geral de proteção de dados comentada. São Paulo: Revista dos Tribunais, 2022.

MARCHAND-MELSOM, Alexander; NGUYEN MAI, Duong Bao. (2020) Automatic repair of OWASP Top 10 security vulnerabilities: A survey. In: Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops. 2020. p. 23-30.

MCCULLOCH, Warren S.; PITTS, Walter. (1943) A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, v. 5, p. 115-133, 1943.

MCKINNEY, Wes. (2022) Python for data analysis. " O'Reilly Media, Inc.", 2022.

MENEZES, P. M., Cardoso, L. M., & Rocha, F. G. (2015). Segurança em redes de computadores uma visão sobre o processo de Pentest. Interfaces Científicas-Exatas e Tecnológicas, 1(2), 85-96. 2015

MORIN, Benjamin et al. (2009) A logic-based model to support alert correlation in intrusion detection. Information Fusion, v. 10, n. 4, p. 285-299, 2009.

OCDE (2013) (Organização para a Cooperação e Desenvolvimento Econômico – OCDE). The OECD privacy framework. 2013.

O'KANE, Philip; SEZER, Sakir; CARLIN, Domhnall. (2018) Evolution of ransomware. Iet Networks, v. 7, n. 5, p. 321-327, 2018.

NIST (2023) Disponível em: https://www.nist.gov/cyberframework. Acessado em 10/08/2023.

PANDYA, Deven; PATEL, N. J. (2016) OWASP top 10 vulnerability analyses in government websites. International Journal of Enterprise Computing and Business Systems, v. 6, n. 1, 2016.

PINHEIRO, Patrícia Peck (2020) Proteção de dados pessoais: Comentários à lei n. 13.709/2018-lgpd. Saraiva Educação SA, 2020.

RAJESH, P. et al. (2022) Analysis of cyber threat detection and emulation using mitre attack framework. In: 2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA). IEEE, 2022. p. 4-12.

SHAH, Sugandh; MEHTRE, Babu M. (2014) An overview of vulnerability assessment and penetration testing techniques. Journal of Computer Virology and Hacking Techniques, v. 11, p. 27-49, 2015. SHEN, Lei. The NIST cybersecurity framework: Overview and potential impacts. Scitech Lawyer, v. 10, n. 4, p. 16, 2014.

SHINDE, Pramila P.; SHAH, Seema. (2018) A review of machine learning and deep learning applications. In: 2018 Fourth international conference on computing communication control and automation (ICCUBEA). IEEE, 2018. p. 1-6.

SILOWASH, George J.; KING, Christopher. (2013) Insider threat control: Understanding data loss prevention (DLP) and detection by correlating events from multiple sources. Softw. Eng. Inst., Carnegie Mellon Univ., Pittsburgh, PA, USA, Rep. CMU/SEI-2013-TN-002, 2013.

SIKORSKI, Michael; HONIG, Andrew. (2012) Practical malware analysis: the hands-on guide to dissecting malicious software. no starch press, 2012.

SINGH, Jagsir; SINGH, Jaswinder. (2018) Challenge of malware analysis: malware obfuscation techniques. International Journal of Information Security Science, v. 7, n. 3, p. 100-110, 2018.

SCHALLER, Valentin. (2022) Combine DLP and EDRM to enable automatized and secure information sharing. 2022. Tese de Doutorado. Alpen-Adria-Universität Klagenfurt.

STROM, Blake E. et al. (2018) Mitre att&ck: Design and philosophy. In: Technical report. The MITRE Corporation, 2018.

ANEXOS

ANEXO I: Professional Guidance of the DPOs – BR in Corporate Governance in Logistics Chains

Ano	Artigo	Status
2022	Professional Guidance of the DPOs – BR in Corporate Governance in Logistics Chains	Liliam Sayuri Sakamoto, Jair Minoro Abe, Jonatas Santos de Souza, Nilson Amado de Souza, Aparecido Carlos Duarte, Edvania Tarkiainem e Luigi Pavarini de Lima.
	Artigo Aceito e Publicado no site	https://link.springer.com/chapter/10.1007/978-3-031-16411-8_8
	SiteScore	1.4
	Highst Percentile	28% 101/140 Information Systems and Management

Professional Guidance of the DPOs-BR in Corporate Governance in Logistics Chains

Liliam Sayuri Sakamoto[1(✉)], Jair Minoro Abe[1], Jonatas Santos de Souza[1], Nilson Amado de Souza[1], Aparecido Carlos Duarte[1], Edvania Tarkiainem[2], and Luigi Pavarini de Lima[3]

[1] Paulista University, 1212, Dr. Bacelar Street, São Paulo, SP, Brazil
liliasakamoto@gmail.com
[2] IMF Smart Education, 25, Bernardino Obregón Street, Madrid, España
[3] Institute of Mathematics and Statistics (IME) - University of São Paulo - USP, 1010, Matão Street, São Paulo, SP, Brazil

Abstract. Currently, DPO (Data Protection Officer) professionals are working in sectors of the economy in the adaptation of the LGPD (Brazilian General Data Protection Law) in Brazil with the use of best management practices. Companies with consolidated corporate governance admit the incorporation of the LGPD (Brazil's General Data Protection Law) into their strategy, with highly trained professionals to consolidate leadership and monitor results. The commitment to enhance the structures that serve logistics companies is preferably used only by highly trained professionals, including internationally. In this sense, it was sought to list points that the sector should develop throughout the adaptation and especially in a continuous way to help achieve the objective within the institution. Data provided and collected in 2022 by the ANPPD (National Association of Data Privacy Professionals) from the almost 4 thousand (associates) trained professionals working in the Brazilian market show us that service sectors (50%) are using professionals trained in the LGPD, as well as in commerce (17%), in the industry (15%), in other segments (14%), and even showed in particular growth in agribusiness (4%) of professionals both in awareness and adaptation of the LGPD. And as strengths that must be followed in both awareness and adequacy, are the continuous improvement in acculturation processes, creation of orientation guides, training, until its implementation by outsourcing (DPO-as-a-Service), and professional responsibilities trained in LGPD and recognized by the CBO (Brazilian Classification of Occupations).

Keywords: Brazilian general data protection law · Protection · Logistics · Corporate governance · Professional guidance

1 Introduction

Identifying the core recognition aspect and its risks sectors in internal and external controls in corporate governance is one of the improvement challenges to DPO – Data

Published by Springer Nature Switzerland AG 2022
D. Y. Kim et al. (Eds.): APMS 2022, IFIP AICT 664, pp. 57–65, 2022.
https://doi.org/10.1007/978-3-031-16411-8_8

Protection Officer in the enterprises in Brazil in compliance with LGPD (Brazil's General Data Protection Law) [3]. The external segments can be converted into better observance throughout Society, with a focus on Corporate Social Responsibility.

In Brazil, advances achieved by the capital market in Brazilian companies began with the Law n° 6.404/1976 and the creation in 1976 of the Brazilian Securities and Exchange Commission, like SEC - Securities and Exchange Commission in the USA. In 1995 with the Brazilian Institute of Corporate Governance (IBGC), and later in the 1988 Constitution [4] and with the national environmental policy. Brazilian business sectors are using the best market practices through Corporate Governance of Information and Technology with the support of specialists recognized by the Brazilian Classification of Occupations [5]. This study aims to elicit the main sectors of the Brazilian economy that are acting in the awareness, adequacy of LGPD and motivators of regulations as well as guidance by experts in LGPD in supply chain too.

1.1 Corporate Governance

Corporate Governance is due to the fact of mitigating misconduct that occurs within the Companies. The path pursued by the implementation and continuous improvement is based on its level of quality (transparency in business conduct, trained professional), and adequacy to regulation. And so, the guarantee of shareholders' rights is allowed [6].

Researchers point out that Corporate Governance allied to Corporate Social Responsibility and Information Technology (IT) brings significant gains for all those involved (Company Departments, Shareholders, Society). Therefore, when implemented by trained managers [9] internationally to work in the public and private sectors. Corporate Governance allied to Information Technology contributes to the improvement of Brazilian companies, as it allows providing all necessary support for strategic definitions under the responsibility of top management. In other countries, this contribution creates many governance innovations for assertive decisions.

COBIT 2019 is a Framework with a structure of Governance and Management internal control objectives [10], it is divided into two parts that address the corporate aspects that are Governance and Management. This guideline helps integrate standards, other guidelines, regulations, and best practices like Six Sigma, ITIL, PMBOOK, Sarbanes-Oxley regulations and community contribution.

Each part is organized into domains, in the Governance part there is only 1 domain and in the Management part, there are 4 domains. In EDM - Evaluate, Direct, and Monitor - Governance domain, there are 5 objectives: Ensured Governance Framework Setting and Maintenance, Ensured Benefits Delivery, Ensured Risk Optimization, Ensured Resource Optimization, and Ensured Stakeholder Engagement [10].

In Management 4 domains: APO - Align, Plan and Organize, BAI - Build, Acquire and Implement, DSS - Deliver, Service and Support, and MEA - Monitor, Evaluate and Assess, there are 35 control objectives. Inside its control objective, there is a description, and this purpose is described as a waterfall model, beginning with Enterprise Goals and this corporate alignment, with example metrics for each enterprise goal match to other example metrics for alignment goals [10].

It Control objective is organized by processes or tasks with organizational structures (Information Flows and Items) and presents the detail of people, skills, competencies, policies, procedures, culture, ethics and behavior, services, infrastructure, and applications. The innovation in this COBIT 2019 version is the addition of Design Factors and Focus Area to tailor the framework to priority objectives, adapt the guidance from the specific focus area, and target the capability and performance customized for each organization, see Fig. 1 [10].

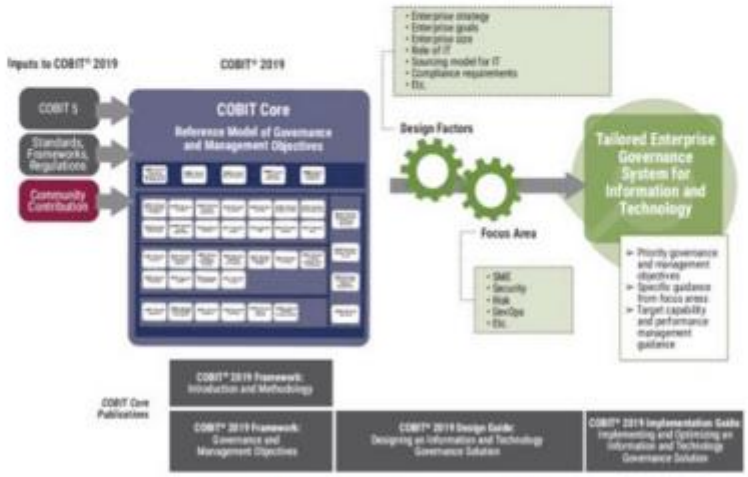


Fig. 1. COBIT 2019 framework [10]

IT Governance Models have been highlighted with the use of frameworks (Model of the Center for Information Systems Research (CISR) of MIT, COBIT 2019 - Control Objectives for Information and related Technology of ISACA- Information Systems Audit and Control Association, Corporate Governance of Information and Technology) as a means of support [10], with a focus aligned with the Company's business. In COBIT 2019, the corporate objectives in item 03 (EG-Enterprise Goals) stand out for dealing with the financial part and its compliance with external laws and regulations.

In COBIT® 2019, one of the models [10] used by Brazilian companies with an emphasis on the design factor, corporate objectives (EG-Enterprise Goals), and especially in the 19 risk categories shown in Table 1 [10]. These risks must be addressed by the DPO, which will be one of the main governance actors.

Table 1. Risk category COBIT® 2019.

Item	Risk category
1	Making IT investment decisions
2	Program and project lifecycle management
3	IT costs and oversight
4	IT expertise, skills, and behaviors
5	Enterprise IT Architecture
6	IT operational infrastructure incidents
7	Unauthorized actions
8	Software adoption and usage issues
9	Hardware incidents
10	Software flaws
11	Logical attacks
12	Third party and supplier incidents
13	Incompatibility
14	Geopolitical issues
15	Industrial action
16	Action of nature
17	Technology-based innovation
18	Environmental issues
19	Data and information management

1.2 Supply Chain Management and Logistics

A Supply Chain Management (SCM) is characterized by the standardization of operations integrated with several processes of purchase/sale, transport, storage, and distribution. This supply chain deals with flows and information [2] and is subject to compliance with the LGPD [3] because sensitive data travel between internal and external processes.

One of the bottlenecks in Brazil in terms of logistics costs is in agribusiness due to the production volume and distances between those who produce and their destinations. The sector in constant evolution seeks improvements in its processes with a Corporate Governance in the adaptation of commercial partners, suppliers, and farmers to adequate social-environmental practices.

Agribusiness evolves and updates itself, as does the awareness [16] of LGPD by trained DPO professionals [15] and its importance [17]. New processes are attributed to the use of the Artificial Intelligence technique [14] for prediction as support, and analysis [18]. Faced with these technological advances in the agribusiness sector, there has been a significant improvement in decision-making.

The adequacy of the LGPD [3], by producers still treated with costs, move towards the awareness of understanding as an investment, since agribusiness deals from confidential

information of employees, suppliers, and customers to the culture of using paper to record personal data that are subject to LGPD [3].

Precision agriculture is also a point of attention as they are susceptible to data leaks and the possibility of having their entire operation interrupted. In this sense, as a response, the sector incorporates Corporate Governance allied to Information Technology.

In systemic terms, a strong ally in the supply chain is the ERP Enterprise Resource Planning tool for logistics [11] that allows numerous advantages, such as real-time information, identification of bottlenecks in processes, better inventory control, reduction in delivery time, compliance with the LGPD [3].

Professional DPOs guide the use of ERP systems approved by the association [1] with methodological quality [7] in their evaluation, as it has been a differential in standardization, and data integration as support in the adequacy by the small company.

And another differentiator for the DPO in the supply chain is knowing about KGM - knowledge governance mechanisms [13] using agri-food supply chain (AFSC) [8].

1.3 LGPD Regulation

Amends the Federal Constitution [11] to include the protection of personal data among fundamental rights and guarantees and to establish the exclusive competence of Brazil to legislate on the protection and processing of personal data.

The regulation of the Law in Brazil has been used as a complement to the understanding of the Law. It is worth remembering that no regulation takes away the right of the Law, as there are always unfounded attempts by parts of regulatory authorities.

We have recently managed to advance and regulate small businesses: Resolution CD/ANPD [12]. Approves the Regulation implementing Law No. 13.709, of August 14, 2018, General Law for the Protection of Personal Data (LGPD) [3], for small treatment agents (micro-companies, small companies, startups, legal entities governed by private law, including non-profits organizations).

In terms of regulations and ordinances, we are at the beginning because we have a long way to go on several points such as the definition of the Anonymization technique[16], portability standard, legacy base sharing, international agreements for data transfer, revocation in the punishment of DPO, among others in its Art.41, §2, III [3], violations in the Brazilian electoral context, and including the performance of inspection by the recently created National Data Protection Authority (ANPD).

Main Motivator for Regulation of the LGPD [3]. Much is researched on the real reasons that lead companies to adhere to adequacy after the awareness stage. The result of a discussion on the topic "Motivator for Regulation" is elicited (see Fig. 2) and shows that the main one is the fear of Fines (36%), followed by Pressure from Customers (26%), Pressure from Customers Regulators (24%), Employee Pressure (10%), International Pressure (2%), Others (2%).

Fig. 2. Main motivator for regulation of LGPD in Companies. (Author).

2 Methodology

The methodology used was the literature review of Corporate Governance, LGPD Regulation e Supply Chain Management and Logistics. It was allied to an analysis of exploratory statistical research conducted by ANPPD on the training of professionals working with data protection and privacy, such as DPOs, mainly working in the supply chain area.

3 Professional Guidance for DPOs

3.1 Data Protection Officer Occupation

In 2016 in Europe, the European Regulation on Privacy of Personal Data [8] was created and served as a model in 2018 for Brazil. This model was the basis for the creation of the General Law on Personal Data Privacy-LGPD [3]. The purpose of the Law is to protect the fundamental rights of freedom and privacy and the free development of the personality of the natural person.

The LGPD – General Personal Data Privacy Law [3] was authorized in 2018, sanctioned by the then President of the Republic Michel Temer on August 14, 2018, taking full force in 2020.

The laws agree that for companies to adapt, it is necessary to involve areas such as IT/Information Security and regulation. And for that, they bring as a requirement a professional responsible for the privacy of personal data, the Data Privacy Officer in Brazil, known worldwide as DPO – Data Protection Officer.

And by LGPD Law in its Section II - Person in Charge of the Processing of Personal Data, its Art.41, §2, III [3] the professional is responsible for guiding and acting in the activities: accept complaints, receive communications, guide employees, and perform other duties. Important and recognized functions [5] by the Brazilian Classification of Occupations.

The Brazilian Classification of Occupations – CBO [5], established by ministerial decree no. 397, of October 9, 2002, aims to identify occupations in the labor market, for classification purposes with administrative and household records:

- CBO Code: 1421-35 – Personal Data Protection Officer.

3.2 Profissional Training

And as strengths that must be followed in awareness and adequacy, they range from continuous improvement in acculturation, and guidance guides, to their implementation by outsourcing and responsibilities of the professional training in the LGPD [3].

These points are the major contribution of DPOs in Brazil as recognized by the CBO [5] (Brazilian Classification of Occupations) body. Due to the recognition of the profession, there is also a growing demand in the service format (DPO as a Service) in hiring a data privacy specialist to fill the role, ensuring compliance with a fundamental point of the General Data Protection Law [16].

About Sector with DPO, Brazilian companies shown in Table 2 reflect the use of IT support (Information Technology). The professionalism of the DPO has been a benchmark among the sectors. The service sector has come out ahead with 50% of the market, just as agribusiness already appears in the survey as initially active with DPO.

Table 2. IT Governance Sectors - 2022 (Author).

Item	Sector	Representation
1st	Services	50%
2nd	Business	17%
3rd	Industry	15%
4th	Others	14%
5th	Agribusiness	4%

4 Conclusion

This research study presented training on data privacy controls to DPO professionals in controllers to help reduce the penalty of fines (36%).

Another point raised in this semester of 2022 was the emergence of agribusiness (4%) in the search for awareness and adequacy of the LGPD through trained professionals.

The data showed that lessons learned by the Corporate Governance of Information and Technology are standing out in the service sector (50%) in the pursuit of LGPD.

Acknowledgements. We thank the research group Paraconsistent logic and artificial intelligence maintained by the Paulista University and conducted by researcher Dr. Abe. This study was financed in part by the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) -Financial Code 001.

References

1. ANPPD – Associação Nacional dos Profissionais de Privacidade de Dados. https://anppd.org/noticia/anppd-homologa-seu-1-software-para-conformidade-a-lgpd-09-09-2021. Accessed 11 Feb 2021
2. Ballou, R.H.: Gerenciamento da cadeia de Suprimento/Logística Empre-sarial, 5th edn. Porto Alegre (2006)
3. Brasil. Lei Geral de Proteção de Dados Pessoais (LGPD). Lei nº 13.709, de 14 de agosto de 2018. http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Accessed 09 Feb 2022
4. Brasil. Constituição Federal, de 1988. http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Accessed 09 Feb 2022
5. CBO - Classificação Brasileira de Ocupações - 5.1.7 (mtecbo.gov.br). https://anppd.org/noticia/cargo-de-dpo-agora-tem-cbo-03-08-2021. Accessed 1 Feb 2022
6. Crisóstomo, V.L., Brandão, I.F.: The ultimate controlling owner and corporate governance in Brazil. Corp. Gov. Int. J. Bus. Soc. **19**(1), 120–140 (2019)
7. de Alencar Nääs, I., et al.: Lameness prediction in broiler chicken using a machine learning technique. Inf. Process. Agric. **8**(3), pp. 409–418 (2021). ISSN 2214-3173. https://doi.org/10.1016/j.inpa.2020.10.003. https://www.sciencedirect.com/science/article/pii/S2214317320302092
8. de Lima, L.A., Abe, J.M., Martinez, A.A.G., de Frederico, A.C., Nakamatsu, K., Santos, J.: Process and subprocess studies to implement the paraconsistent artificial neural networks for decision-making. In: Jain, V., Patnaik, S., Popenţiu Vlădicescu, F., Sethi, I.K. (eds.) Recent Trends in Intelligent Computing, Communication and Devices. AISC, vol. 1006, pp. 503–512. Springer, Singapore (2020). https://doi.org/10.1007/978-981-13-9406-5_61. 2019 Print ISBN 978-981-13-9405-8. ISBN 978-981-13-9406-5
9. de Souza, J.S., Abe, J.M., de Lima, L.A., de Souza, N.A.: The general law principles for protection the personal data and their importance. In: 7th International Conference on Computer Science, Engineering, and Information Technology (CSEIT 2020) (2020). Computer Science & Information Technology (CS & IT), Copenhagen, Denmark, Anais 2020. vol. 10, p. 109. https://doi.org/10.5121/CSIT.2020.101110. https://arxiv.org/abs/2009.14313
10. de Souza, J.S., et al.: The Brazilian law on personal data protection. Int. J. Net. Secur. Appl. (IJNSA) **12** (2020). SSRN https://ssrn.com/abstract=3949175
11. Dora, M., et al.: Critical success factors influencing artificial intelli-gence adoption in food supply chains. Int. J. Prod. Res., 1–20 (2021)
12. European Comission GDPR - General Data Protection Regulation (2016). https://gdpr-info.eu/. Accessed 11 Feb 2022
13. Gangi, F., Meles, A., Monferrà, S., Mustilli, M.: Does corporate social responsibility help the survivorship of SMEs and large firms? Glob. Finance J. **43**, 100402 (2020)

14. ISACA: Information Systems Audit and Control Association – COBIT 2019: Introduction and Methodology. ISACA, Rolling Meadows (2019b)
15. Lima, L.A., et al.: DPO no Brasil sob a ótica da LGPD – Lei Geral de Proteção de Dados. Instituto EXIN-Ministry of Economic Affairs in The Netherlands (2020). https://www.exin.com/br-pt/dpo-no-brasil-sob-a-otica-da-lgpd-lei-de-protecao-dedados/
16. PEC 17/2019- Proposta de Emenda à Constituição 1988. Available in: https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2210757. Accessed on: 10/02/2022
17. RESOLUTION CD/ANPD No. 2, 27 January 2022. https://www.in.gov.br/en/web/dou/-/resolucao-cd/anpd-n-2-de-27-de-janeiro-de-2022-376562019#:~:text=A%20ANPD%20poder%C3%A1%20determinar%20ao,os%20riscos%20para%20os%20titulares. Accessed 10 Feb 2022
18. Zhao, G., et al.: The impact of knowledge governance mecha-nisms on supply chain performance: empirical evidence from the agri-food industry. Prod. Plan. Control **32**(15), 1313–1336 (2021)

ANEXO II: Overview of DPOs in the use of LGPD compliance software with Logic Paraconsistent Annotated Evidential Eτ and DLP – Data Loss Prevention

Ano	Artigo	Status
2022	**OVERVIEW OF DPOS IN THE USE OF LGPD COMPLIANCE SOFTWARE WITH EVIDENTIAL ANNOTATED PARACONSISTENT LOGIC Eτ AND DLP – DATA LOSS PREVENTION**	Liliam Sayuri Sakamoto, Jair Minoro Abe, Aparecido Carlos Duarte e José Rodrigo Cabral
	Aceito, apresentado e publicado, ISBN	978-625-8246-28-5
	Publicado no Site	https://en.iksadkongre.net/kongre-kitaplari
	SiteScore	Não consta
	Highest Percentile	Não consta

CUKUROVA 9th INTERNATIONAL SCIENTIFIC RESEARCHES CONFERENCE
9 - 11 October 2022 / Adana, TURKEY

OVERVIEW OF DPOS IN THE USE OF LGPD COMPLIANCE SOFTWARE WITH EVIDENTIAL ANNOTATED PARACONSISTENT LOGIC Eτ AND DLP – DATA LOSS PREVENTION

Msc. Liliam Sayuri SAKAMOTO
Paulista University, 1212 Dr. Bacelar Street, São Paulo, Brazil.
ORCID: 0000-0001-8636-0100
Dr. Jair Minoro ABE
Paulista University, 1212 Dr. Bacelar Street, São Paulo, Brazil.
ORCID: 0000-0003-2088-9065
Aparecido Carlos DUARTE
Paulista University, 1212 Dr. Bacelar Street, São Paulo, Brazil.
ORCID: 0000-0002-3278-3471
José Rodrigo CABRAL
Paulista University, 1212 Dr. Bacelar Street, São Paulo, Brazil.
ORCID: 0000-0003-3896-6233

ABSTRACT

A DPO – Data Protection Officer is a professional who needs to be always up to date with local legislation regarding data protection and privacy.

There is a concern with the implementation of effective models both in Brazil and in the world, but most of the time the action of a DPO is based on mitigating controls, however at times, due to a large amount of information that must be analyzed, he will need to of tools such as software for compliance with the LGPD – General Data Protection Law of Brazil.

This study addresses exploratory research on this software that considers the use of Paraconsistent Evidential Annotated Logic Eτ, to direct the contradictions in decision making combined with the practice of DLP - Data Loss Prevention.

These results are also compared with the statistics of the ANPPD - National Association of Data Privacy Professionals, which has about 4,000 professionals, presented at the National Conference of Data Privacy Professionals of the CNPPD held two times: in March/2021 and March/2022, presenting a comparative overview of DPOs in Brazil.

Keywords: LGPD – General Data Protection Law; Evidential Annotated Paraconsistent Logic Eτ; DPO – Data Protection Officer, DLP – Data Loss Prevention

INTRODUCTION

This study presents a comparative professional overview between 2021 and 2022 of the DPO - Data Protection Officer in compliance with the LGPD and their performance in Brazilian companies [1].

The GDPR - General Protection Data Regulation in the European Union presents specific documentation for DPO compliance and professional duties [14], while in the LGPD it is described only in article 41[1].

Therefore, for the DPO to be able to validate compliance with the LGPD, it cannot always do this manually and will need the help of automated tools [11]. Software for LGPD compliance is not always approved under the terms of the legislation, and the ANPPD Scientific Committee performs an analysis of this compliance based on classical logic and the project to implement an assertiveness of this study for an annotated paraconsistent logic Eτ [3].

Exploratory research was carried out to improve the approval of software in compliance with the LGPD, using paraconsistent logic, to optimize the decision-making of the evaluating DPO [5].

This study presents a literature review on the LGPD Law specifically on article 41 which details the role of the DPO. The methodology used in the applied exploratory research is the Annotated Evidential Paraconsistent Logic Eτ, to balance the existing contradictions in the approval of the software for LGPD. The statistics collected by the ANPPD - National Association of Data Privacy Professionals at the CNPPD National Conference of Data Privacy Professionals in March/2021 and in March/2022 will also be compared [3][5].

1.1. LGPD - General Data Protection Law of Brazil

With the approval of the LGPD in 2020 in Brazil, there was a need for companies to adapt to the Law, although many of the companies are still not fully compliant with the legislation [1].

The LGPD deals with access to information and the authorization of use, to guarantee the privacy of such data only to the legal person, that is, the data subject [2].

The performance of the DPO within the business areas can be observed in the periods 2021 and 2022, as shown in fig. 1:

Exploratory research was carried out to improve the approval of software in compliance with the LGPD, using paraconsistent logic, to optimize the decision-making of the evaluating DPO [5].

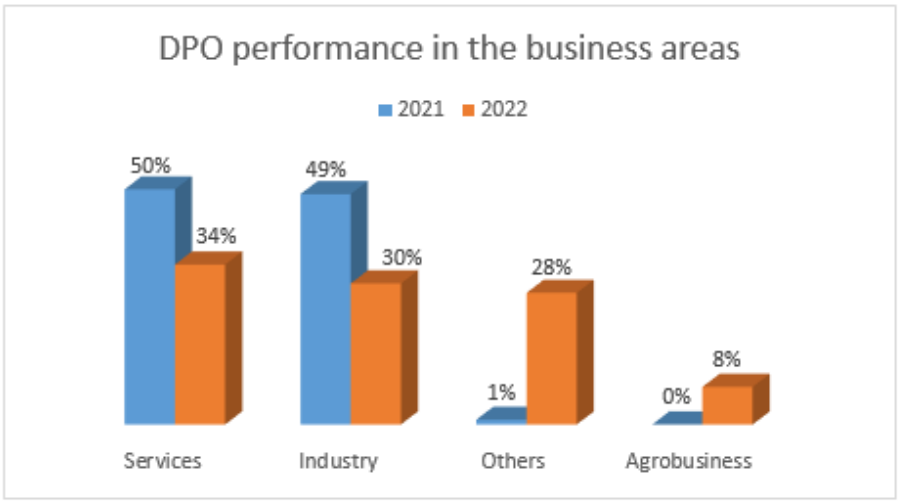
This study presents a literature review on the LGPD Law specifically on article 41 which details the role of the DPO. The methodology used in the applied exploratory research is the Annotated Evidential Paraconsistent Logic Eτ, to balance the existing contradictions in the approval of the software for LGPD. The statistics collected by the ANPPD - National Association of Data Privacy Professionals at the CNPPD National Conference of Data Privacy Professionals in March/2021 and in March/2022 will also be compared [3][5].

1.1. LGPD - General Data Protection Law of Brazil

With the approval of the LGPD in 2020 in Brazil, there was a need for companies to adapt to the Law, although many of the companies are still not fully compliant with the legislation [1]. The LGPD deals with access to information and the authorization of use, to guarantee the privacy of such data only to the legal person, that is, the data subject [2].

The performance of the DPO within the business areas can be observed in the periods 2021 and 2022, as shown in fig. 1:

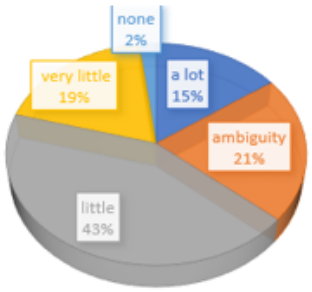
(a) Fig. 1:



It is observed that from one period to the next, there was an increase in other business areas, in addition to industry and services, with agribusiness being the highlight.

The awareness of companies regarding the need to adhere to the LGPD presented a situation of attention in 2021, however, compliance to adhere to legislation in 2022 can be observed, as shown in fig. 2:

(a)



(b)

Fig.2: (a) 2021 (b)2022

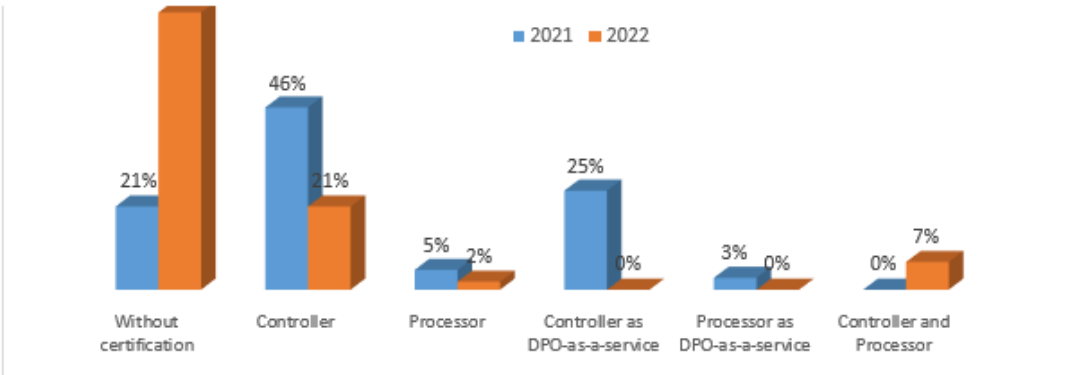
1.2. DPO – Data Protection Officer function

For both the LGPD and the GDPR, the role of the DPO - Data Protection Officer is indicated to orchestrate the company's compliance with the legislation, in addition to being the focal point in the interaction with data subjects and with the supervisory body. The DPO may be a Controller employee or a Processor for this specific task.

To safely adhere to the LGPD, the professional DPO - Data Protection Officer, must be certified to perform this function. Some certifying institutions are internationally recognized, such as EXIN and IAPP.

The ANPPD - National Association of Data Privacy Professionals is a class association that brings together certified professionals in Brazil and makes their surveys available free of charge on its website [5]. In fig. 3, in 2021 there were not so many DPOs working without certification, now in 2022, a large part of the professionals in the function do not have certification or background, a situation that concerns the adherence to LGPD in an assertive way.

Fig. 3



1.3. Paraconsistent Evidential Annotated Logic Eτ

Logic presents a language of the propositional annotated logics Pτ, and the atomic propositions are of the type $p_{(\mu, \lambda)}$ where p is a proposition and μ, λ ∈ [0, 1] (closed unit real interval), μ indicates the degree of favourable evidence1) of p and λ the degree of contrary evidence of p. Annotated logics are based on a lattice of truth-values denoted by τ [6]. The values μ, λ depending on the considered applications and may change in fact, μ maybe the degree of belief favourable and λ maybe the degree of contrary belief of the proposition p; also, μ can indicate the probability expressed by p to occur and λ the improbability expressed by p of occurring. The atomic propositions $p_{(\mu, \lambda)}$ of logic Eτ can be read as: I believe in p with the degree of favourable belief μ and the degree of unfavourable belief λ, or the degree of favourable evidence of p is μ and the degree of unfavourable evidence of p is λ{6].

A proposition p(μ, λ) can be read as: " The favorable evidence of p is μ and the unfavorable evidence is λ" [7]. For instance, p (1.0, 0.0) can be read as a true proposition, p (0.0,1.0) as false, p (1.0, 1.0) as inconsistent, p (0.0, 0.0) as paracomplete, and p (0.5, 0.5) as an indefinite proposition [8]. Also we introduce the following concepts: Uncertainty degree: Gun (μ, λ) = μ + λ - 1 (0 ≤ μ, λ ≤ 1) and Certainty degree: Gce (μ, λ) = μ - λ (0 ≤ μ, λ ≤ 1) [9].

An order relation is defined on [0, 1]: (μ1, λ1) ≤ (μ2, λ2) ↔ μ1 ≤ μ2 and λ 2 ≤ λ1, constituting a lattice that will be symbolized by τ.

Table 1. Extreme states[6].

Extreme States	Symbol
True	V
False	F
Inconsistent	T
Paracomplete	⊥

Table 2. Extreme states[6].

Non-Extreme States	Symbol
Near-true tending to Inconsistent	QV→T
Quasi-true tending to Paracomplete	QV→⊥
Quasi-false tending to Inconsistent	QF→T
Quasi-false tending to Paracomplete	QF→⊥
Almost-inconsistent tending to True	QT→V
Almost-inconsistent tending to False	QT→F
Quasi-paracomplete tending to True	Q⊥→V
Quasi-paracomplete tending to False	Q⊥→F

Four outer limit values [6]:

Vcve = C1 = Veracity control value; 0 ≤ Vcve ≤ 1

Vcfa = C2 = False control value; -1 ≤ Vcfa ≤ 0

Vcic = C3 = Inconsistency control value; $0 \leq$ Vcc ≤ 1

Vcpa = C4 = Paracompleteness control value; $-1 \leq$ Vcpa ≤ 0

Such values will direct when a proposition is considered, for example, "true" in the sense that we make a positive decision, and so on.

(a)

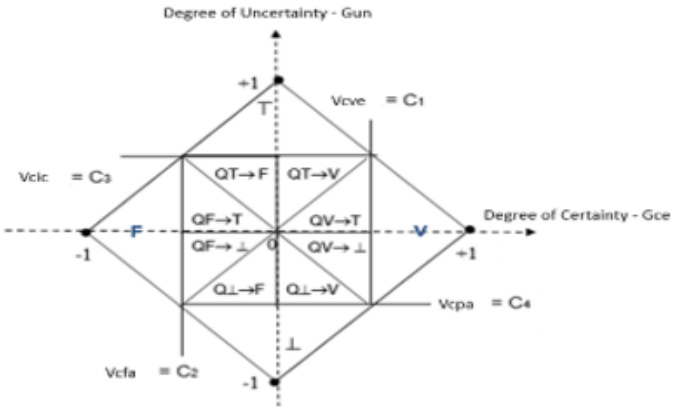


Fig.4: (a) Extreme and Non-extreme states that represent table 2 [6]

1.4. DLP – Data Loss Prevention

A DPO needs to use a combination of several analysis techniques and tools to assist him in his role as the main actor to direct compliance with LGPD or other legislation in the corporation where he operates and may be subordinate to a Controller or a Processor. For this reason, this study addresses one of the points where most companies have a gap, which is the leakage of sensitive and personal data or information, which is the most critical asset that is manipulated today [15].

And the use of tools with Artificial Intelligence such as those that can be configured called DLP - Data Loss Prevention, sometimes only added to some antivirus or some monitoring of internal traffic on the network, but they have a limitation in their assertiveness. Some just scan files or emails looking for just a few keywords, like the credit card number [15].

1.5. Software compliance with LGPD

Despite the creation of several types of market software for compliance with the LGPD to assist the DPO in its work for validation with the legislation, many are not approved or validated according to the regulation [5].

The ANPPD Scientific Committee certifies and approves LGPD-compliant software that can be used by large and medium-sized companies. There are 4 approval phases to LGPD – compliance software, table 3.

This study analyzed the software approval process of the Scientific Committee of the ANPPD that uses classical logic and suggests its optimization with the use of Paraconsistent Logic allied to DLP - Data Loss Prevention.

Table 3. Approval phases to LGPD – compliance software.

Phase	Description (*a*)	LGPD compliance(b)	DLP Compliance (c)
1	Intuitive	Operacional Leve	Operational Data
2	Defined	Internal Controls	Technical Data
3	Measurable	Process Monitoring	Differentiations of Corporate Data and Personal Data
4	Optimized	Decision Making based on indicators	Strategic Data

For software to be approved in this process, there are some maturity criteria from one phase to another. However, here we will only be analyzing the transition from the software that is in Phase 1 to Phase 2, table 4.

2. METHODOLOGY

First, a literature review was carried out, in which the most detailed points of the LGPD Law were addressed, especially those that directly link the DPO's function [14]. It presented how the approval of the software was used for adequacy and compliance with the LGPD [1].

This approval process originated from a work by the Scientific Committee of the ANPPD, in which the authors are active members, and their analysis of statistics carried out in its Conference [3] [5] led to the hypothesis about the weighting regarding the use of Paraconsistent Logic to optimize this software approval process since today it is carried out based only on classical logic.

The methodology used was applied exploratory research, observing the process of homologation of the software for LGPD compliance with 4 (four) phases. For this study, only phase 2 validation was initially chosen.

The purpose of this study is to present the consideration on the optimization of the software approval process for data protection and privacy prepared by the Scientific Committee of the ANPPD with the use of Paraconsistent Logic instead of classical logic to align the contradictory ideas of experts of the area, combined with the DLP – Data Loss Prevention technique, where there is a mapping of the information that can be leaked, as shown in Table 3.

For that, 2 preponderant factors named, such as: "F1" and "F2" were selected. For Factors 1 and 2, only 1 section will be addressed, which will be called "S1" for both, although the content is different, the formation will be the same. These Factors and Sections were chosen according to their adherence to the software approval process, as well as with the LGPD. They are detailed in table 5 below:

Table 5. Factors and Sections Source

Factors	Sections
F1 – General overview about LGPD	S1 – Is there motivation in companies for adopting LGPD Act?
F2 – Software maturity stages	S1 - Are there formalized of Data Protection and Privacy Practice? (Phase 2)

Data collection was carried out through an online questionnaire to capture the quantitative data described in tables 4 and 5. Eighteen professionals responded, where the mathematical average of the values found was calculated and grouped into:

• Group 1: three Processors and three steps.
• Group 2: six DPOs - Data Protection Officers as a Service.
• Group 3: six Controllers.

The groups were chosen based on the most important functions of the LGPD [1].

The analysis of statistics carried out at the 2021[5] and 2022[3] ANPPD Conference was also the object of this applied research.

In the construction of the database, values from 0 to 1 were collected, in which, through an online survey carried out with the three groups of selected specialists, detailed in tables 6 and 7.

Table 6. Authors: Factors and Sections Source

FACTORS	SECTIONS	Group 1 -Processors						Group 2 - DPO´s						Group 3 - Controllers					
		Expert 1		Expert 2		Expert 3		Expert 4		Expert 5		Expert 6		Expert 7		Expert 8		Expert 9	
		μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ
F1	S1	0,9	0,1	0,9	0,1	0,9	0,3	0,9	0,2	1,0	0,3	0,8	0,1	0,9	0,3	1,0	0,1	1,0	0,2
F2	S1	0,2	1,0	0,3	1,0	0,3	0,9	0,4	0,6	0,3	0,7	0,2	0,9	0,1	0,8	0,2	0,9	0,3	0,9

Table 7. Authors: Factors and Sections Source

FACTORS	SECTIONS	Group 4 - Lawyers						Group 5 - DPO´s						Group 6 - Controllers					
		Expert 10		Expert 11		Expert 12		Expert 13		Expert 14		Expert 15		Expert 16		Expert 17		Expert 18	
		μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ
F1	S1	0,7	0,1	0,9	0,3	0,8	0,1	0,9	0,1	1,0	0,1	1,0	0,3	0,8	0,1	0,7	0,1	1,0	0,1
F2	S1	0,1	1,0	0,1	0,9	0,1	0,9	0,3	0,8	0,2	0,7	0,1	0,9	0,2	0,8	0,2	0,9	0,2	1,0

The results of Tables 6 and 7:

Table 8. Authors: Average over μ and λ factors and sections

Result	Factor	Section	Conclusion		Definition
			μ	λ	
A	F1	S1	0,9	0,1	True
B	F2	S1	0,3	0,9	Inconsistent

3. DISCUSSION AND RESULTS

The results were applied in Para-analyser Algorithm, and we obtained:

(a)

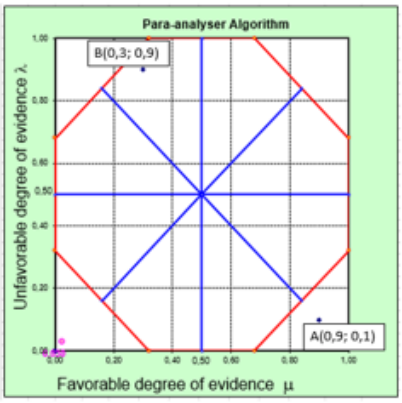
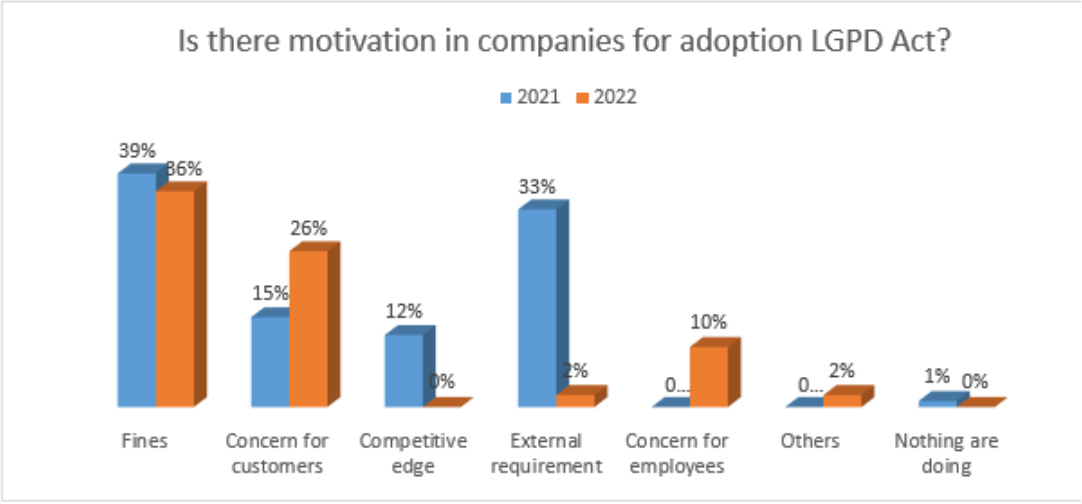


Fig. 5: (a) Results in Para-analyser algorithm

The results of table 8 in comparison with the CNPPD statistics in 2021 and 2022 lead us to the following conclusions:

About "A" – Factor 1 – Section 1: Is there motivation in companies to adopt the LGPD Act?
The result is true and analyzed in Fig. 6 the presented statistics observed that:
(a) 2021: 99% motivation for adopting LGPD, so only 1% is doing nothing.
(b) 2022: 100% motivation for adopting LGPD.
In this case, there lined up.

(a)



About "B" – Factor 2 – Section 1: Are there formalized Data Protection and Privacy practices? (Phase 2) The result is inconsistent analyzed. In this case, the software used in LGPD validation is not prepared to increase to level 2 of maturity. At this point, it is possible to implement the DLP – Data Loss Prevention for earlier Identification and later refine it with the use of Paraconsistent Evidential Annotated Logic Eτ, as shown in fig. 7.

(a)

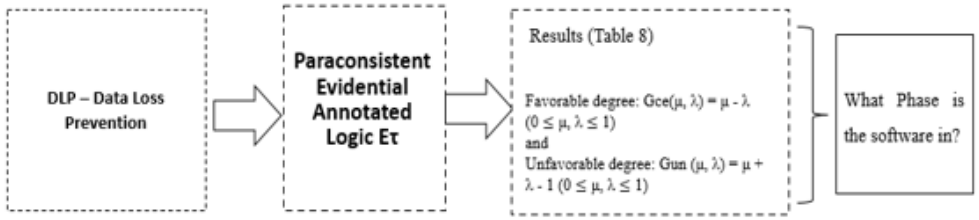


Fig. 7: DLP – Data Loss Prevention + Paraconsistent Evidential Annotated Logic Eτ

4. CONCLUSION

All authors are required to complete the Procedia exclusive license transfer agreement before the article can be published, which they can do online. This transfer agreement enables Elsevier to protect the copyrighted material.

Many companies to remain in compliance with the LGPD need specialized professionals to act in this direction the DPOs.
For the performance of these professional DPOs to be assertive, this study identified with the exploratory research applied using Paraconsistent Evidential Annotated Logic Eτ that Brazilian companies are effectively seeking to adapt to the regulation of the LGPD.
And what increasingly targeted tools or a set of processes and techniques can also optimize the DPO's performance to ensure compliance with the Law, such as the combination of DLP - Data Loss Prevention and Paraconsistent Evidential Annotated Logic Eτ can be a differential for refinement and assertiveness in the decision-making.

ACKNOWLEDGMENTS

We thank the research group Paraconsistent logic and artificial intelligence maintained by Paulista University and conducted by researcher Dr. Abe. This study was financed in part by the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financial Code 001.

REFERENCES

[1] Brasil. Lei Geral de Proteção de Dados Pessoais (LGPD). Lei nº 13.709, de 14 de agosto de 2018. Available in :http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm . Accessed on: 21/04/2022.

[3] Lima, Luiz A. Comitê Científico ANPPD Estatísticas: Panorama da Conscientização Nacional sobre LGPD. CNPPD 2022. Available in : https://cnppd.online/. Accessed on: 21/04/2022.

[4] Bioni, Bruno Ricardo. Proteção de Dados Pessoais: a função e os limites do consentimento. Ed 1. Vol. único. Rio de Janeiro: Foresee, 2019

[5] Lima, Luiz A. Comitê Científico ANPPD Estatísticas: Panorama da Conscientização Nacional sobre LGPD. CNPPD 2021. Available in : https://cnppd.online/. Accessed on: 21/04/2022.

[6] Abe, J. M. Nakamatsu K. Introduction to Annotated Logics - Foundations for Paracomplete and Paraconsistent Reasoning, Series Title Intelligent Systems Reference Li brary, Volume 88, Publisher Springer International Publishing, Copyright Holder Springer International Publishing Switzerland, eBook ISBN 978-3-319-17912-4, DOI 10.1007/978-3-319-17912-4, Hardcover ISBN 978-3-319-17911-7, Series ISSN 1868-4394, Edition Number 1, 190 pages, 2015.

[7] De Carvalho, F.R., Abe, J.M.: Tomadas de decisão com ferramentas da Lógica Paraconsistente Anotada. São Paulo. Blucher, pp. 37–47, 2011.

[8] Abe, J.M., et al.: Lógica Paraconsistente Anotada Evidencial Et, pp. 38–39. Comunnicar, Santos, 2011.

[9] De Carvalho, F.R., Brunstein, I., Abe, J. M.: Paraconsistent Annotated Logic in Analysis of Viability: in approach to product launching. In: Dubois, D.M. (ed.), vol. 718, pp. 282–291, 2011.

[10] Dill, R.P., Da Costa Jr., N., Santos, A. A. P.: Corporate Profitability Analysis: A Novel Application for Paraconsistent Logic. Applied Mathematical Sciences 8, 2014.

[11] De Lima L.A., Abe J.M., Martinez A.A.G., De Frederico A.C., Nakamatsu K., Santos J. (2020) Process And Subprocess Studies To Implement The Paraconsistent Artificial Neural Networks For Decision-Making. In: Jain V., Patnaik S., Popențiu Vlădicescu F., Sethi I. (Eds) Recent Trends In Intelligent Computing, Communication And Devices. Advances In Intelligent Systems And Computing, Vol 1006. Springer, Singapore. 2019 Print ISBN: 978-981- 13-9405-8; Online Isbn: 978-981-13-9406-5; Available in: https://doi.org/10.1007/978-981-13-9406-5_615

[12] European Comission Guidelines on Data Protection Officers ('DPOs') (wp243rev.01) 2016 Available in :<https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612048> Accessed on: 21/04/2022.

[13] European Comission Guidelines on Consent under Regulation 2016/679 (wp259rev.01) 2016 Available in :<https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051 > Accessed on: 21/04/2022.

[14] European Comission GDPR - General Data Protection Regulation 2016 Available in :< https://gdpr-info.eu/ > Accessed on: 21/04/2022.

[15] SILOWASH, George J.; KING, Christopher. Insider threat control: Understanding data loss prevention (DLP) and detection by correlating events from multiple sources. CARNEGIE-MELLON UNIV PITTSBURGH PA SOFTWARE ENGINEERING INST, 2013.[16]

ANEXO III: Optimizing the Data Loss level using Logic Paraconsistent Annotated Evidential Eτ

Ano	Artigo	Status
2023	Optimizing the Data Loss level using Paraconsistent Annotated Logic Eτ	Liliam Sayuri Sakamoto, Jair Minoro Abe, Jonatas Santos de Souza, Luiz Antonio de Lima.
	Publicado como Capitulo do Livro: Advances in Applied Logics (Intelligent Systems Reference Library book series)	https://link.springer.com/chapter/10.1007/978-3-031-35759-6_9
	SiteScore	2.0
	Highest Percentile	62% 99/266 Library and Information Sciences

Optimizing the Data Loss level using Paraconsistent Annotated Logic Eτ

Liliam Sayuri Sakamoto[1[0000-0001-8636-0100]], Jair Minoro Abe[2[0000-0003-2088-9065]], Jonatas Santos de Souza[3[0000-0002-0052-0132]], Luiz Antonio de Lima[4[0000-0003-4228-2387]]

[1,2,3,4,5] Paulista University, 1212 Dr. Bacelar Street, São Paulo, Brazil.
liliam.sakamoto@aluno.unip.br

Abstract. Currently, corporations worldwide have a problem that grows exponentially: orchestrating the organization and understanding structured and unstructured data. These unstructured data can be grouped in repositories, for isolated and random data entry, however on the other hand the data loss analysis, that is, data loss prevention, where some criteria of artificial intelligence create templates that are monitored, and that, because they are very restricted, also present contradictions and flaws. The focus of this study is to optimize this analysis by minimizing the level of data loss using the Paraconsistent Evidential Annotated Logic Eτ. With a bibliographic review on DLP - Data Loss Prevention, Paraconsistent Evidential Annotated Logic Eτ, Artificial Intelligence techniques, and data protection. With the use of a Python program, and applied research will be carried out with data from a financial company, which presents 40% of data loss in its analysis with this process of Artificial Intelligence, in comparison with the minimization of this data loss with the use of Evidential Annotated Paraconsistent Logic Eτ to 25%, that is, a difference of 15%.

Keywords: DLP – Data Loss Prevention, Paraconsistent Logic Evidential Annotated Eτ, Artificial Intelligence.

1. Introduction

1.1 General context

It is interesting how with the evolution of operational and production processes, there was also an evolution in capturing data and transforming them into feasible information. Today, digitally structured data can be easily identified in a company, as they are within the essential systems and in organized repositories. However, how to identify all other unstructured data, both physical, isolated, and merged in social networks, text messages, emails, cloud drives, saved on flash drives or external hard drives, stored on the end users' hard drives in their endpoints, and the data that is transiting from one point to the other?[11]

The differential of each company to identify this data is in its risk appetite because the more the data and the transformed information become an asset of increasing value, the greater the importance of its security and organization. Understanding the flow of these input, transformation, output, and transit data brings security that there is active monitoring and that data loss can be analyzed through Artificial Intelligence algorithms.

Despite this type of tool being advanced in contrast to the manual analysis of data, some data may present doubts, causing problems in identifying the monitored data template. For this reason, the implementation of another analysis layer with the Annotated Paraconsistent Logic Evidential Eτ adds greater assertiveness in decision making.

A sample of a financial company that asks for confidentiality was collected, noting that the data were anonymized in compliance with the LGPD - General Data Protection Law of Brazil to test this implementation [22], only referring to a period of 1 month of data.

1.2 General Data Protection Law of Brazil

The protection of personal data [21], public or private, sensitive or not, is directly related to the protection of the intimacy and private life of individuals. It is worth remembering that the right to privacy is, as a rule, born with a negative aspect, that is, the right not to be molested.

According to the LGPD [22], it is the citizen's right to know, correct, and determine to whom, how, when, and how their data will be disclosed. It can be seen that the revolution brought about by the advent and diffusion of the Internet has given a new meaning to the right to privacy.

Therefore, the citizen has the right to be alone and the possibility of demanding concrete benefits, that is, demanding information, correcting it, and controlling its use. It is interesting to note that the need for legal protection of the citizen originates in the realization that the personal data that circulate on the web have an economic content, that is, that there is the possibility of commercialization of such data.

The impact of this Law, in line with technological evolution and the need for business innovation brought about by this new scenario, led directly to creating economic, social, and political effects in the country. As a result, there is a need for professionals with specialized knowledge in specific topics aimed at protecting personal data, which is called in the GDPR of the DPO (Data Protection Officer) and in the LGPD of the person in charge of personal data.

DPO – Data Protection Officer will have the challenge of coordinating and bringing companies into compliance, not only with the LGPD but with other laws that deal with the matter, such as the Consumer Protection Code, Marco Civil da Internet Law [22], Federal Constitution[2]; in addition to the laws, there are also information security standards (normative standards, among others) [22] that should guide the security practices applied. As if that were not enough, they must still know how to talk about business because aiming at all this compliance, it is necessary to map processes and understand the flow of information, people involved, and technologies to support the entire process.

A new law was recently enacted in Brazil, arousing the population's collective interest, governmental and non-governmental bodies, and business segments. The General Law for the Protection of Personal Data, Law no. 13,709 [23] of August 14, 2018, grants the Brazilian population rights and guarantees on how organizations will adapt to the collection and processing of personal data, whether by physical or digital means.

Companies are now concerned with protecting people's data due to the approval of the Brazilian Law that was named Marco Civil da Internet [23].

Brazil leads in America in terms of uniqueness in terms of awareness, regulation, new businesses, as it brings together internationally trained professionals and officially recognized on July 29, 2021, as the occupation of DPO/Data Processing Officer was included in the Brazilian Classification of Occupations (CBO) which will begin in 2022.

The Data Officer has a vital role in this mission provided for in Art. 41, and in order not to run the risk of suffering from the high fines that the LGPD will charge anyone who is not in compliance with its rules, the ideal is to find a professional to help the company in the transition.

The relevance of Data Privacy has greatly increased in recent years. The theme gained more evidence after the scandals involving social media companies accused of selling personal data to other marketing companies, there was also the case with the email of former secretary of defense Hilary Clinton, among others. This triggered serious postures in several countries in relation to information security and the result was the creation of strict laws on the subject. In Brazil, DPO professionals unite through the APDADOS – National Association of Data Privacy Professionals in Brazil.

A great example is the GDPR, the European Regulation on Privacy of Personal Data that served as a model for the Brazilian law LGPD – General Personal Data Privacy Law authorized in 2018, sanctioned by the then President of the Republic Michel Temer (PMDB) on December 14. August 2018, with full force as of 2020.

Both laws agree that, in order to adapt to companies, it is necessary to involve areas such as IT/Information Security and Legislation; for this, they bring as a requirement a professional responsible for the privacy of personal data, the Data Privacy Officer in Brazil, known worldwide as DPO – Data Protection Officer.

During this scenario, several information security professionals and lawyers began to observe this new field of activity; however, unlike lawyers, IT professionals did not have an organization that represented the class in the decisions that were being processed in the National Congress regulating the details about the role of the DPO.

Even in 2019, before the emergence of the ANPD - National Data Privacy Authority, there was discussion in Brazil whether the person's figure in charge would only fit for lawyers or IT professionals. It was then that in June 2019, MP 869/2018 had in its text the addition of the term "legal-regulatory" as a prerequisite of knowledge for a professional to exercise the role of DPO; because of this, more questions were asked, and such decision would benefit only the legal class.

Seeing this movement, some academic executives and enthusiasts who were part of the technical committees in Brasília raised the flag that it was time for IT professionals to have a representation focused on the topic of privacy and data privacy. At that moment, such a group of academic executives was looking for names that could start a visionary project of national magnitude like this, without political/partisan bias, and aftermarket research, they pointed out the name of Dr Davis Alves to chair such an initiative.

Davis Alves is an IT professional, PhD in the United States, specializing in information security with several international certifications and being one of the first DPOs in Brazil to work abroad, having trained the first Brazilian Data Protection Officers.

In the first half of 2019, Dr Davis Alves accepted the challenge and started to gather people interested in the subject, having his students from the EXIN Privacy & Data Protection Practitioner course from Portal do Training as enthusiasts and future DPOs, who joined as

founding members to form the then APDADOS – Associação Brazilian Data Privacy Professionals After the initial meeting with the founding members, Dr Davis Alves sought out big names in the IT area to join the APDADOS steering committee, among them:

Umberto Correia, DPO - IT governance and information security executive of one of the largest Brazilian institutions for the vice presidency;

André Masili, DPO and founder of Grupo Linx/SA for the general secretary; The APDADOS was founded with the mission of bringing together the best Data Privacy Professionals – DPOs in Brazil, promoting technical/scientific knowledge on the subject and representing the class during decision-making in the National Congress involving the LGPD by supporting it. Those with technical bases without partisan and political ends.

In the preliminary provisions of the LGPD, art. 6 reinforces the activities of processing personal data, by the principles of good faith, where information security must use technical and administrative measures capable of protecting this data from unauthorized access and from incidents that cause loss, alteration, communication or dissemination[1].

Another relevant aspect presented by LGPD in its article 11 is the processing of sensitive personal data, which are those that can discriminate against a person, such as racial or ethnic origin and religious conviction, which can only be manipulated in case of prevention of fraud, security of data holder, identification process and registration authentication in electronic systems, exceptions provided for in art. 9 where there is a need to guarantee the protection of this holder's data [1].

The LGPD values data security and confidentiality according to art.46, where processing agents and processors must use security, technical and administrative measures capable of protecting access to unauthorized personal data and from accidental or illegal incidents, such as loss, alteration, and communication, or inappropriate treatment[2].

1.3 Artificial Intelligence

Research related to AI started after World War II, and Alan Turing carried out the first work in this area. Since then, much research has been carried out. Defining the concept of artificial intelligence is very difficult. For this reason, Artificial Intelligence was (and remains) a notion that has multiple interpretations, often conflicting or circular.

The difficulty of a clear definition may come from the fact that several human faculties are being reproduced, from the ability to play chess or be involved in computer vision, voice analysis and synthesis, fuzzy logic, artificial neural networks, and many others. Initially, AI aimed to reproduce human thought. Artificial

Intelligence embraced the idea of reproducing human faculties such as creativity, self-improvement, and by use of language Artificial Neural Networks.

The Recurrent Neural Network (RNN) network architecture: The hidden neurons of the recurrent neural network receive the result of the mathematical operation that they performed in the previous time and the data from the previous layer. Because they have this characteristic, these networks can model problems with temporal characteristics, such as the weather forecast given the climate history in a past window. Thus, the RNNs consider a temporal dependency between the input data.

1.4 Machine learning

Neural networks can "learn" to diagnose the most appropriate antibiotic, but they depend on preliminary tests with a structured database evaluated by specialists in actual cases. There are several methods for training the neural network.

The activation function receives the sum of the multiplication of each weight by the respective value of the parameter (or sign) evaluated; this is typically done by the scalar product of the weight matrix by the parameter value matrix. The result of this sum serves as an input parameter for an activation function where a trigger threshold is defined. If the result of the calculation of the activation function exceeds the threshold, a value is propagated to the next neuron ahead. The nonlinear activation function usually gives better results. As well in figure 4:

Activation function Sigmoid σ (z): (ranging from 0 to +1): activation notation used in linear function. For each large z, we have σ (z): = 1; for z, small or tending to small, we have σ (z) ≈0. In training, the average sigmoid is around 0.5. They are usually used only in the Output layer when a binary is expected.

Tanh-hyperbolic activation function (ranging from -1 to +1): activation notation used in non-linear function. It is generally used in the output and hidden (learning) layer because in the it is close to 0.0.

Activation function Rectified Linear Unit ReLU: activation notation used in non-linear function. Usually used in the hidden layer (learning).

Activation (look at figure 1) function Rectified Linear Unit ReLU-Leaky [Loose] (max): activation notation used in non-linear function. Usually used in the hidden layer (learning). we have σ (z): = 0.01;

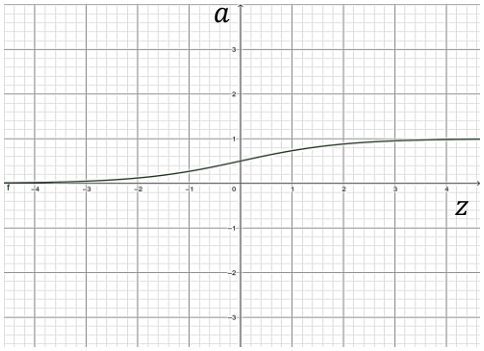


Figure 1 Sigmoid$a = \frac{1}{1 + e^{-z}}$

2. Bibliographic Review

2.1 DLP – Data Loss Prevention

Data Loss Prevention is not new, but it is being used more nowadays, as ready-made tools on the market already have some features of this monitoring standard built-in. Applications such as Office 365 are already able to bring some templates internally standardized for compliance with the GDPR. Even trying to do an analysis for compliance with the regulations, with the lack of depth in the evaluation of certain criteria, data loss can occur [15].

Other Antivirus tools like Trend Micro, which has the solution called Deep Security for servers only, and another solution called Apex One, which is a SaaS for endpoint security, both have the functionality of reporting DLP logs – Data Loss Prevention, which can be configured according to the company's needs, for example: search for information disclosure of CPF, CNPJ, Invoices, credit card, social security, etc. [15].

Internally, each of these tools works with artificial intelligence algorithms that cross data and monitor patterns pre-formatted by the information security analyst, but directly using classical logic for their conclusions and report presentations so that a manager can make a decision-

making process. The Decision whether to continue with the process when a data loss situation is detected.

DLP – Data Loss Prevention helps to delimit the data:

• what are their origins or inputs?

• where they are stored.

• within which systems undergo consistency and are transformed into informative reports.

• what are the data outputs?

• and when, how, and where they can be destroyed

Our study delimits the data source into structured and unstructured, where the data that is in:

• From the Network (on the servers and application systems).

• From the Cloud (external storage).

• From Storage (local storage).

• From the endpoints (end-user storage).

While the unstructured ones are:

- Printed reports, such as old reports from decommissioned legacy systems.
- Data on microfilm or microfiche, technology discontinued.
- Data on social networks such as WhatsApp, Instagram, and Facebook.
- Data on USB sticks and external hard drives.
- Data in cloud drives such as OneDrive, SharePoint, Teams, Google drive, etc.

The DLP needs an assertive configuration by a specialist analyst in the area to capture possible results with the company's needs, such as credit card purchase patterns, which was one of the areas that first developed standards for monitoring suspicious situations [16].

Another widely used situation for DLP standardization is for suspected malware, ransomware, and spam, as some indicators can be copied from public safety lists called CVEs.

Not all Incident detection tools can offer their customers very competitive optimization services; however, many need a very refined and carefully protected management, usually by a SIEM - Security Information and Event Monitoring, that is, a tool that manages to unify alerts about malicious activities such as IP scanning, data flow, malicious emails, intrusion attempts, firewall monitoring, review of security policy updates, security patches implemented, but here we will use this concept to create a specific process focused on Data Loss[16].

Additionally, better opportunities to effectively identify redundant and non-ideal elements in organizational and management structures that can be safely identified, changed, or removed, bringing the benefit of pre-emption before the data management framework [15].

If there is a demand regarding the outdated business environment, it brings awareness about the need to carry out a complete optimization of the organization and data privacy security management structure preferably using aspects of DLP - Data Loss Prevention to prevent loss or data leakage [17].

The tools usually need some testing and calibration time, so they do not create alerts with excessive false positives but use classic logic [19].

The level of sophistication reached by malicious software requires constant efforts and considerable expenditure of resources to mitigate this practice. The balance of risk appetite between the parties is fragile, as recognized by several authorities in the area. Therefore, one must constantly innovate and be one step ahead of these offenders [18].

2.2 Lógica Paraconsistente Anotada Evidencial Eτ

Annotated logics constitute a class of paraconsistent logic. knowledge of an expert on an analyzed subject, questions are used to capture opinions that are normalized in logic between 0 and 1, as shown in Figure 3.These values are respectively the favourable evidence that is expressed by the symbol μ and the contrary evidence by λ.

Logic Eτ must follow the process (look at figure 2) during the application, which can be seen in the figure:

The definition means understanding and creating the proposition to be understood and which should reflect the problem.

The transformation would be treating data from favourable and unfavourable evidence. We usually normalize the answer between the number zero and one to be handled by Logic Eτ.

Thus, it becomes possible to perform the processing of calculations of the data collected. With this, we have the Favorable Degree and the Degree of Unfavorable.

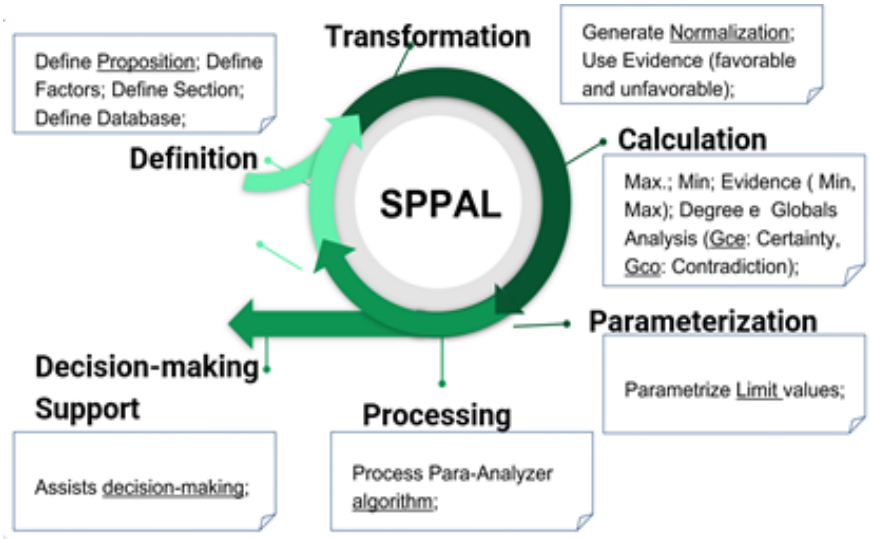


Figure 2 Steps Process Paraconsistent Annotated Logic

The acceptable limits are parameterized to obtain an analysis that makes the data and can be transformed into useful information.

The intelligence is given when applying the para-analyzer algorithm that contains all the information to execute the Eτ Logic.

As summarized by Abe et al. [8], programs can now be built using paraconsistent logic, making it possible to treat inconsistencies directly and elegantly. This feature can be applied in specialist systems, object-oriented databases, representation of contradictory knowledge, etc., with all the implications in artificial intelligence.

In Abe et al. [8]: "Evidential Annotated Paraconsistent Logic Eτ has an Eτ language and the atomic propositions are of the type p (μ, λ) where p is a proposition and μ, λ∈ [0, 1]. Intuitively, μ indicates the degree of unfavourable evidence1 of p and λ the degree of contrary evidence of p. The reading of the values μ, λ depends on the applications considered and may change: in fact μ it may be the degree of belief favourable and λ it may be the degree of belief contrary to proposition p; also, μ can indicate the probability expressed by p occurring and λ the improbability expressed by p occurring. The atomic propositions p (μ, λ) of logic Eτ can intuitively be read as: I believe in p with the degree of favourable belief μ and the degree of contrary belief λ, or the degree of favourable evidence of p is μ and the degree of evidence to the contrary of p is λ".

Paraconsistent logics can serve as underlying logic of theories in which A and ¬A (the negation of A) are both true without being trivial [6]. There are many types of paraconsistent systems. In this text, it consider the Paraconsistent Annotated Evidential Logic Eτ. The formulation in Logic Eτ are of the type p (μ, λ), in which p is a proposition and e (μ, λ) ϵ [0, 1] is the real unitary closed interval.

A proposition p(μ, λ) can be read as: " The favorable evidence of p is μ and the unfavorable evidence is λ" [8]. For instance, p (1.0, 0.0) can be read as a true proposition, p (0.0,1.0) as false, p (1.0, 1.0) as inconsistent, p (0.0, 0.0) as paracomplete, and p (0.5, 0.5) as an indefinite proposition [8]. Also we introduce the following concepts: Uncertainty degree: $G_{un\,(\mu,\,\lambda)} = \mu + \lambda - 1$ $(0 \leq \mu, \lambda \leq 1)$ and Certainty degree: $G_{ce(\mu,\,\lambda)} = \mu - \lambda$ $(0 \leq \mu, \lambda \leq 1)$ [9].

An order relation is defined on [0, 1]: $(\mu_1, \lambda_1) \leq (\mu_2, \lambda_2) \leftrightarrow \mu 1 \leq \mu 2$ and $\lambda\, 2 \leq \lambda 1$, constituting a lattice that will be symbolized by τ.

With the uncertainty and certainty degrees, we can get the following 12 output states (Table 2): extreme and non-extreme states. It is worth observed that this division can be modified according to each application [20].

Para-analyser Algoritm.

In this proposed algorithm, there is a set of information obtained, which can sometimes seem contradictory, making it difficult to analyze the scenario for risk analysis. Generally, in such situations, this information is discarded or ignored, that is, they are considered "dirty" of the system, however at best they may even receive different treatment.

Silva Filho, Abe and Torres [7] quote: "However, the contradiction most of the time contains decisive information, as it is like the encounter of two strands of opposing truth values. Therefore, to neglect it is to proceed in an anachronistic way, and that is why we must look for languages that can live with the contradiction without disturbing the other information. As for uncertainty, we must think of a language that can capture the 'maximum' of 'information' of the concept".

In this line of reasoning for the analysis based on Paraconsistent Logic, situations of Inconsistency and Paracompleteness will be considered together with the True and False, represented according to Table 1:

Table 1 – Extreme States Fonte: Abe et al. (2011).

Extreme States	**Symbol**
True	V
False	F
Inconsistent	T
Paracomplete	⊥

The set of these states or objects ($\tau = \{F, V, T, \perp\}$) can also be called annotation constants and can be represented using the Hasse diagram as shown in figure 3:

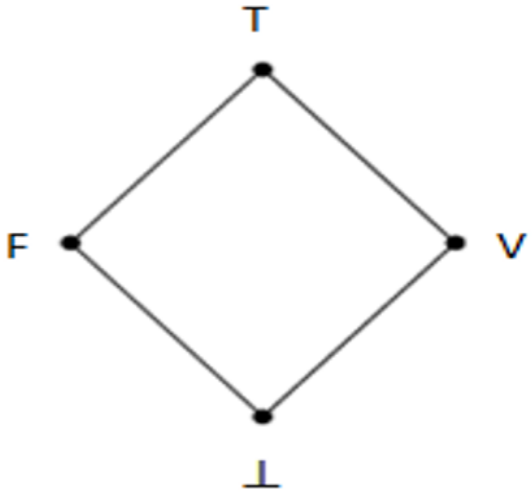


Figure 3 Lattice diagram

"The operator about τ é: ~:|τ| →|τ| that will operate, intuitively, like this:
~ T = T (the 'negation' of an inconsistent proposition is inconsistent)
~ V = F (the 'negation' of a true proposition is false)
~ F = V (the 'negation' of a false proposition is true)
~⊥ = ⊥ (the 'negation' of a paracomplete proposition is paracomplete)
Annotated Paraconsistent Logic will be used; this type must be composed of 1, 2 or "n" values.

With the calculations of the values of the axes that make up the representative figure of the lattice, it can be divided or internally delimited into several regions of different sizes and formats, thus obtaining a discretization of the same. From the bounded regions of the lattice, it is possible to relate the resulting logical states, which, in turn, will be obtained by interpolating the Degrees of Certainty Gc and Contradiction Gct. Thus, for each interpolation between the degrees of certainty and contradiction, it is possible to extract information to assist in decision making [20].

The representation of table 2 shows a representation of the lattice constructed with values of Degrees of Certainty and Contradiction and sectioned into 12 states. Thus, at the end of the analysis, one of the 12 possible resulting logical states will be obtained as an answer for decision making.

Table 2. Non-extreme states

Non-extreme States	Symbol
Quasi-true tending to Inconsistent	QV⟶T
Quasi-true tending to Paracomplete	QV⟶⊥
Quasi-false tending to Inconsistent	QF⟶T
Quasi-false tending to Paracomplete	QF⟶⊥
Quasi-inconsistent tending to True	QT⟶V
Quasi-inconsistent tending to False	QT⟶ F
Quasi-paracomplete tending to True	Q⊥⟶V
Quasi-paracomplete tending to False	Q⊥⟶ F

Some additional control values are:

- V_{scct} = maximum value of uncertainty control = Ft_{un}
- V_{scc} = maximum value of certainty control = Ft_{ce}
- V_{icct} = minimum value of uncertainty control = $-Ft_{un}$
- V_{icc} = minimum value of certainty control = $-Ft_{ce}$

All states are represented in the next figure 4.

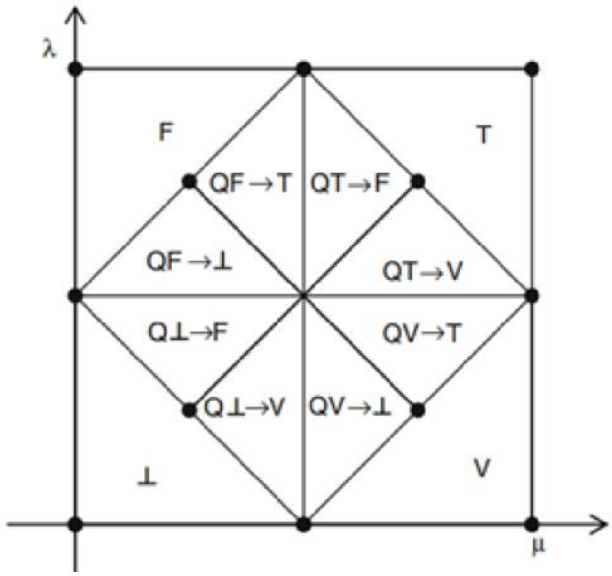


Figure 4 Aspect of the Lattice to make decision Source: (Abe et al., 2011).

2.3 Artificial Intelligence Techniques

In the 1940s, Warren McCulloch and Walter Pitts made the first proposition of an artificial intelligence model, as shown in Fig. x, which used a neuronal structure, suggesting the use of hardware; in this case, they were variable resistors connected to amplifiers, to behave (look at figure 5). like a human neuron [11].

Its concept was simple, as it was able to model separable linear systems, such as logical operators AND, OR, and NOT. The neuron would input a list of Boolean data (0 or 1), do the sum, and then, in the sequence, do the sum for a triggering function that would return 1 if the sum exceeds the limit and 0 if it fails [11].

Fig. X exemplifies how a neuron would be configured to compute x1 AND !X2. Making !x2 an inhibitory input, where there would be only two possible situations: x1 = 0 and x2 = 0 and x1 = 1 and x2 = 0. Obviously, the expression can only evaluate to be true if x1 = 1, and therefore case 2 is the only valid one [11].

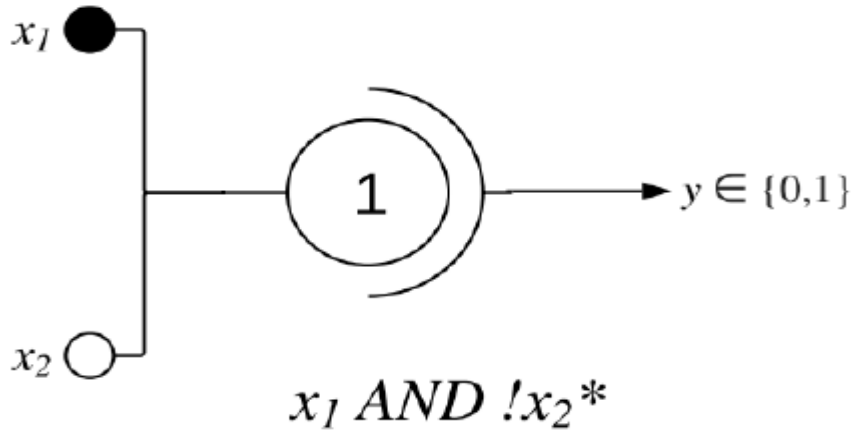


Figure 5. McCulloch-Pitts Neuron Model (1943)

Nowadays, many organizations are looking for AI - Artificial Intelligence solutions to find analogies that lead to optimized containment, being able to use the focus on non-classical logic to circumvent these types of attacks from more offensive groups[20].

The first results of the implementation of methods of convolutional neural networks were carried out with the financial sector, which still prefers not to hire exclusive data analysts because this type of attack attempts that use Artificial Intelligence/Neural Networks, despite having complex problems such as optimization of its cyber structure, according to a study presented by the company Trend Micro. Even the results of the first tests demonstrate the potential of such approaches, which can stimulate companies to think about creating an AI - Artificial Intelligence separate from the internal structure or developing a Machine Learning structure specific to each client[11].

The artificial intelligence technique [12] comes from the 1940s in the studies proposed by Warren McCulloch and Walter Pitts (1943). Some researches are based on neurons' philosophy, knowledge, and function and use propositional logic created by Russell and Whitehead. Another pillar is Turing's theory of computation.

These two researchers proposed a model of artificial neurons, in which each neuron is characterized by being "on" or "off", with the switch to "on" occurring in response the output of a neuron from the layer stimulates the learning of the next layer..

The state of a neuron was considered "equivalent in concrete terms to a proposition that defined its appropriate stimulus". McCulloch and Pitts also suggested that properly defined networks would be able to learn. For example, they showed that a certain network of connected neurons could calculate any computable function and that simple network structures could implement all logical connectives (and, or, not, etc.).

In figure 6, the paraconsistent neuron [1] is proposed to serve the paraconsistent neural network. Another important point must be taken into account that the neuron chooses which characteristic (x) will be used in the network to be trained.

Figure 6 Paraconsistent Artificial Neuron proposal, author, 2021

The concept of Eτ Logic applied [24] in the day-to-day of our reality in the face of numerous sources of information, contradiction constantly occupies a space, bringing uncertainties that will culminate in brief or future challenges.

In the case of a system with artificial intelligence [27], neural networks [27], also known as “machine learning” [28], which starts from the study of pattern recognition [20][21], the appearance of contradiction in reasoning logic is inevitable when we try to reflect human behaviour.

In activities of the medical, hospital, and health segment, it has been observed in the use of analysis of clinical exams, early diagnosis of cancer [26], in politics, in the analysis of lawsuits, and productivity of public security [30], in the measurement of software [25][29] technical support, in the service of insurance companies, where at least two specialists are involved [27], there will always be different points of view.

In response to the contradiction, we have the Eτ Logic in service in any commercial, scientific segment that can be aggregated with other technologies. An application in six sigma services [31] has been explored both in industry and in services where at least two specialists are involved [27].

The Eτ Logic has been researched in the segment of animal welfare; it has always been in-depth with prediction and the entire agribusiness chain [32].

2.4 Data Protection

Brazil has the LGPD - General Data Protection Law in the same standards as the GDPR - General Data Protection Regulation of the European Union; perhaps there is not as much detail as this one and presents the regulatory compliance needs for all Brazilian companies[1][2].

Despite these needs, many companies in Brazil are still not prepared to comply with the LGPD[1], which can be observed in the research carried out by the class entity called APDADOS - National Association for Data Protection Professionals, which annually surveys its 4 thousand members (look at figure 7) across the country and concluded in Fig.1 that[3]:

• 84% of companies are in the implementation process.

• 13% are aware of the GDPR compliance needs but do not apply it.

• 3% do nothing.

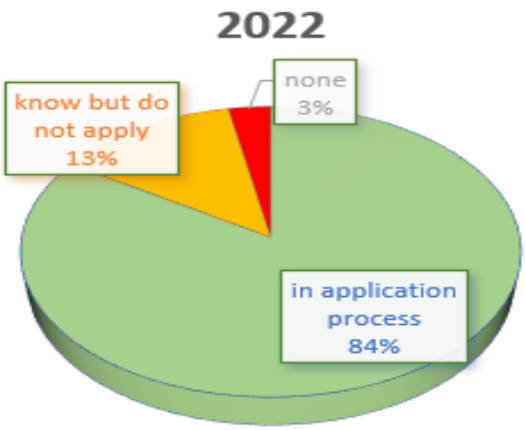


Figure 7. Level of awareness for LGPD compliance in 2022.

Another important point for the development of compliance with the rules is the updating of specialists to mitigate or reduce data loss, this constant concern of both the specialist and the company, as mitigation actions are not only important for adherence to the LGPD. In addition to raising awareness of the entire company together, so that the past problem is not repeated, resulting in fines costs for data leakage incidents [13][14].

Therefore, the analysis of APDADOS statistics [3][5] shows this panorama and indicates the concern with the development of adequate training for professionals working with data protection because the better the level of corporate awareness, the lower the operational error due to ignorance of the risk factor (look at figure 8). In this research, it can be observed that:

- 51% of professionals take specific private courses on LGPD[1].
- 19% participate only in Congresses.
- 16% prefer free courses.
- 14% take a long-term specialization such as an MBA.

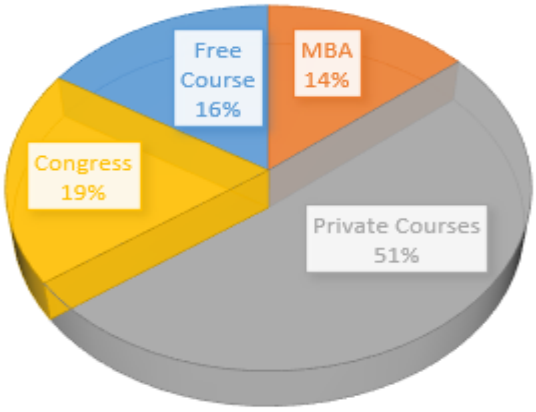


Figure 8. LGPD training

The LGPD details in article 5 the data life cycle, which is determined by the following steps[1]:
- Creation that is determined in: collection, production, reception, or data extraction.
- Transport through transmission, distribution, and communication of data.
- Handling focused on: classification, use and modification of data.
- Storage is determined in archiving and storage.
- Discard when data is deleted.

3 Minimization of data loss

3.1 DLP – Data Loss Prevention using Paraconsistent Evidential Annotated Logic Eτ

When using the concept of DLP – Data Loss Prevention (look at figure 9), the following granularity was considered to obtain graduation for the types of data loss detected [4]:
First level:
• Strategic Data: Loss of confidential data from top-secret projects with greater risk and company exposure in the market.
• Tactical and Technical Data: Loss of confidential information, such as user credentials, system administrators, and authority to transfer large amounts.
• Operational Data: Loss of information on standardization of activities, generating rework or initial mapping needs, and new training for base teams.
The second level [6]:
• Identification of business areas.
• Identification of impacted users.
• Identification of impacted product.
• Identification of incidents with personal data.
• Identification of the cost of the fine due to non-compliance with LGPD
On the third level [11]:

Use the mass of data collected from the financial company for only one month to be analyzed by the Python Algorithm to obtain the report with the answers.

At the fourth level [20], the information from the exceptions presented by the Python Algorithm report is aligned. They are inserted into the Algorithm Para analyzer to identify the percentage of the difference between the responses of the two Algorithms, thus achieving greater assertiveness in making decisions.

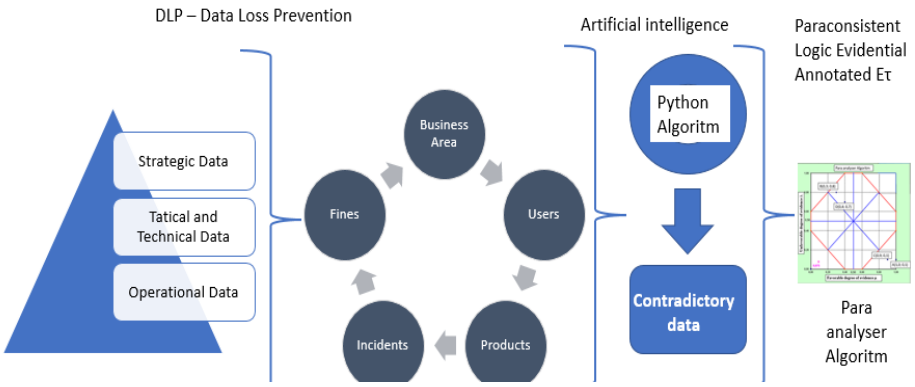


Figure 9. DLP – Data Loss Prevention using Paraconsistent Logic Evidential Annotated Eτ

DLP Tools - on the market have been increasingly sought after by Companies due to LGPD in Brazil, and GDPR in the EU- European Union. Such features can be used as input to Et Logic. Control can be made available via the USB Port like other devices.

The main objective is always to block, monitor, and manage these devices. Granular control based on the identification (ID) of the distributor, supplier, customer, product ID, and serial number are explored by the systems.

There are several channels that can be used including Data Digitization in Motion, making it possible to monitor and block file transfers. All detail is done by inspection of content and context.

Manual or automatic scans are constantly performed to exclude confidential data and ensure DLP. In some cases, data in transit is forcibly encrypted and thus maintains the quality of the DLP technique in data processing.

The Law understands that processing must be considered any operation carried out with personal data, such as those relating to collection, production, reception, classification, use, access, reproduction, transmission, distribution, processing, archiving, storage, elimination, evaluation or information control, modification, communication, transfer, dissemination or extraction

Internationally, data originating from outside the national territory and which are not the subject of communication, shared use of data with Brazilian processing agents, or the subject of international data transfer with another country other than the country of origin, as long as the country of origin provides a degree of protection of personal data appropriate to the provisions of the Law.

The use of DLP tools focuses on complying with regulations regarding sensitive personal data: personal data on racial or ethnic origin, religious conviction, political opinion, membership of a trade union or organization of a religious, philosophical, or political nature, data relating to health or sexual life, genetic or biometric data when linked to a natural person.

Companies must use the DLP concept to process data according to the operator and/or controller precepts. The latter is considered a natural or legal person, governed by public or private law, who is responsible for decisions regarding the processing of personal data. And when an operator is a natural or legal person, under public or private law, who processes personal data on behalf of the controller.

4 Tests

4.1 Python program and mass data results

The mass of data to be tested had the following layout with this legend:

The First Level:

Strategic data= 1.1

Tactical and Technical data=1.2

Operational data=1.3

The second level:

• Identification of business areas=2.1

• Identification of impacted users=2.2

• Identification of impacted product=2.3

• Identification of incidents with personal data=2.4

• Identification of the cost of the fine due to non-compliance with LGPD=2.5

The third level:

effectively-identified data=3.1

contradictory data=3.2

The fourth level:

True=4.1

False=4.2

Incomplete=4.3

Paracomplete=4.4.

For which each (look at figure 10) of the captured lines presented a description of the data source.

	A	B
1	Data	
2	1.1,2.1,3.1,4.1,social-midia	
3	1.1,2.2,3.1,4.2,e-mail	
4	1.1,.2.3,3.2,4.2,storage	
5	1.2,2.4,3.25,4.2,endpoint	
6	1.3,2.5,3.2,4.2,cloud	
7	1.3,2.1,3.1,4.4,cloud	
8	1.3,2.2,3.1,4.3,social-midia	
9	1.3,2.3,3.1,4.2,storage	
10	1.1,2.5,3.1,4.2,endpoint	
11	1.1,2.1,3.1,4.1,social-midia	
12	1.1,2.2,3.1,4.2,e-mail	
13	1.1,.2.3,3.2,4.2,storage	

Figure 10. Data Source example

The mass of data captured presented about 30 pieces of information per day for one month, comprising an amount for analysis of 1,107 data, of which the algorithm verified 60% in Python (look at figure 11) as effective and 40% as contradictory.

Data considered contradictory were passed on to the para-analyzer algorithm, and a differential of 9% true, 6% incomplete, 10% paracomplete and 15% false was obtained.

However, the interval considered false can pass a more excellent sieve that requires more effective monitoring due to the percentage of occurrences that could be an alert of data leakage in a subtle way.

```
import csv
from typing import List
from util import normalize_by_feature_scaling
from network import Network
from random import shuffle

if __name__ == "__main__":
    financ_parameters: List[List[float]] = []
    financ_classifications: List[List[float]] = []
    financ_species: List[str] = []
    with open('\\Users\Liliam\Downloads\RNTreinaCompara\RedeNeural_01\financ.csv', mode='r') as financ_fi
        financses: List = list(csv.reader(financ_file))
        shuffle(financses) # get our lines of data in random order
        for financ in financses:
            parameters: List[float] = [float(n) for n in financ[0:4]]
            financ_parameters.append(parameters)
            species: str = financ[4]
            if species == "social-midia":
                financ_classifications.append([1.0, 0.0, 0.0])
            elif species == "e-mail":
                financ_classifications.append([0.0, 1.0, 0.0])
            else:
                financ_classifications.append([0.0, 0.0, 1.0])
            financ_species.append(species)

    normalize_by_feature_scaling(financ_parameters)

    financ_network: Network = Network([4, 6, 3], 0.3)

    def iris_interpret_output(output: List[float]) -> str:
        if max(output) == output[0]:
            return "social-midia"
        elif max(output) == output[1]:
            return "e-mail"
        else:
            return "storage"
```

Figure 11. Python Program

Conclusion

It is concluded that the alignment between the use of Paraconsistent Evidential Annotated Logic Eτ, with the DLP - Data Loss Prevention presents a significant gain for the issue of monitoring and analysis of data loss.

The implementation of this type of tool can increase the company's performance regarding the level of information security, which despised 40% of the data as inconclusive to only 15%, considering that even these intervals can be monitored so that its effectiveness of 100% of the data transited, kept stored and sent by email or analyzed social networks.

Intelligent tools that support companies to mitigate data loss are seen as strategic in corporations.

Both Brazilian and European Union legislation requires the company to respect the data subject and increase the treatment of data surrounding the constant regulation in new businesses.

It is understood that the union of data privacy professionals with the APDADOS brings the possibility of significant advances in the adequacy of the LGPD and an increase in the flow of new business between consolidated countries within the European Union block using the GDPR.

References

[1] Brasil. Lei Geral de Proteção de Dados Pessoais (LGPD). Lei nº 13.709, de 14 de agosto de 2018. Available in :http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm . Accessed on: 21/04/2022.

[3] Lima, Luiz A. Comitê Científico APDADOS Estatísticas: Panorama da Conscientização Nacional sobre LGPD. CNPPD 2022. Available in : https://cnppd.online/. Accessed on: 21/04/2022.

[4] Bioni, Bruno Ricardo. Proteção de Dados Pessoais: a função e os limites do consentimento. Ed 1. Vol. único. Rio de Janeiro: Foresee, 2019

[5] Lima, Luiz A. Comitê Científico APDADOS Estatísticas: Panorama da Conscientização Nacional sobre LGPD. CNPPD 2021. Available in : https://cnppd.online/. Accessed on: 21/04/2022.

[6] Abe, J. M. Nakamatsu K. Introduction to Annotated Logics - Foundations for Paracomplete and Paraconsistent Reasoning, Series Title Intelligent Systems Reference Library, Volume 88, Publisher Springer International Publishing, Copyright Holder Springer International Publishing Switzerland, eBook ISBN 978-3-319-17912-4, DOI 10.1007/978-3-319-17912-4, Hardcover ISBN 978-3-319-17911-7, Series ISSN 1868-4394, Edition Number 1, 190 pages, 2015.

[7] De Carvalho, F.R., Abe, J.M.: Tomadas de decisão com ferramentas da Lógica Paraconsistente Anotada. São Paulo. Blucher, pp. 37–47, 2011.

[8] Abe, J.M., et al.: Lógica Paraconsistente Anotada Evidencial Et, pp. 38–39. Comunnicar, Santos, 2011.

[9] De Carvalho, F.R., Brunstein, I., Abe, J. M.: Paraconsistent Annotated Logic in Analysis of Viability: in approach to product launching. In: Dubois, D.M. (ed.), vol. 718, pp. 282–291, 2011.

[10] Dill, R.P., Da Costa Jr., N., Santos, A. A. P.: Corporate Profitability Analysis: A Novel Application for Paraconsistent Logic. Applied Mathematical Sciences 8, 2014.

11] De Lima L.A., Abe J.M., Martinez A.A.G., de Frederico A.C., Nakamatsu K., Santos J. "Process and Subprocess Studies to Implement the Paraconsistent Artificial Neural Networks for Decision-Making". In: Jain V., Patnaik S., Popențiu Vlădicescu F., Sethi I. (Eds) Recent Trends in Intelligent Computing, Communication and Devices. Advances in Intelligent Systems nd Computing, Vol 1006. Springer, Singapore. 2019 Print ISBN: 978-981- 13-9405-8; Online Isbn: 978-981-13-9406-5; HTTPS://DOI.ORG/10.1007/978-981-13-9406-5_61

[12] European Comission Guidelines on Data Protection Officers ('DPOs') (wp243rev.01) 2016 Available in :<https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612048> Accessed on: 21/04/2022.

[13] European Comission Guidelines on Consent under Regulation 2016/679 (wp259rev.01) 2016 Available in :<https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051 > Accessed on: 21/04/2022.

[14] European Comission GDPR - General Data Protection Regulation 2016 Available in :< https://gdpr-info.eu/ > Accessed on: 21/04/2022.

[15] SILOWASH, George J.; KING, Christopher. Insider threat control: Understanding data loss prevention (DLP) and detection by correlating events from multiple sources. CARNEGIE-MELLON UNIV PITTSBURGH PA SOFTWARE ENGINEERING INST, 2013.

[16] Priest, Graham, Koji Tanaka, and Zach Weber, "Paraconsistent Logic", The Stanford Encyclopedia of Philosophy (Summer 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2018/entries/logic-paraconsistent/>. ISSN: 1095-5054

[17] Sikorski, M., Honig, A. "Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software", 2012, No Starch Press.

[18] Singh, J., Singh, J., "Challenges of Malware Analysis: Obfuscation Techniques", 2018. Disponível em: http://50.87.218.19/ijiss/index.php/ijiss/article/view/327

[19] Akama, S. Towards Paraconsistent Engineering, Intelligent Systems Reference Library, Germany: Springer (2016).

[20] SUBRAHMANIAN, V.: On the semantics of quantitative logic programs. In: Proceedings of the 4th IEEE Symposium on Logic Programming, pp. 173–182 (1987)

[21] Luiz Antônio de Lima, Liliam Sayuri Sakamoto, Nilson Amado de Souza, Roberto Aures Antonio de Moura, Davis Alves, Claudio Pessoa, José Rogério Poggio Moreira, Jonatas Santos de Souza. "DPO no Brasil sob a ótica da LGPD - Lei Geral de Proteção de Dados". Instituto EXIN - Ministry of Economic Affairs in the Netherlands. 2020. https://www.exin.com/br-pt/dpo-no-brasil-sob-a-otica-da-lgpd-lei-de-protecao-de-dados/

[22] Jonatas S. de Souza, Jair M. Abe, Luiz A. de Lima, Nilson A. de Souza. "The General Law Principles for Protection the Personal Data and their Importance". In: 7th International Conference on Computer Science Engineering and Information Technology (CSEIT 2020 - https://arxiv.org/abs/2009.14313), 2020. Computer science & information technology (cs & it), Copenhagen, Denmark. Anais 2020. V. 10. P. 109. HTTP://DX.DOI.ORG/10.5121/CSIT.2020.101110

[23] JONATAS S. DE SOUZA, JAIR M. ABE, LUIZ A. DE LIMA, AND NILSON A. DE SOUZA, "The Brazilian Law on Personal Data Protection", International Journal of Network Security & Its Applications (IJNSA - https://airccse.org/journal/jnsa20_current.html), November 2019, Volume 12, Number 6, ISSN: 0974-9330[online]; 0975-2307 [Print]. https://aircconline.com/ijnsa/V12N6/12620ijnsa02.pdf

[24] De Lima, Luiz A.; Abe, Jair M.; Kirilo, Caique Z.; Da Silva, Jonas P.; Nakamatsu, Kazumi. "Using Logic Concepts in Software Measurement." PROCEDIA COMPUTER SCIENCE, v. 131, p. 600-607, 2018. http://dx.doi.org/10.1016/j.procs.2018.04.302

[25] Lima, L. A., Abe, J. M., Martinez, a. A. G., Santos, J., Albertini, G., & Nakamatsu, K. (2019). “The Productivity Gains Achieved in Applicability of The Prototype AITOD with Paraconsistent Logic in Support in Decision-Making in Project Remeasurement”. Proceedings of the 9th International Conference of Information and Communication Technology [ICICT-2019] Nanning, Guangxi, China January 11-13, 2019 (http://aivr.org/index.html). Edited by Srikanta Patnaik Volume 154, Pages 1-844 (2019). Procedia Computer Science, pp. 347–353. https://doi.org/10.1016/j.procs.2019.06.050

[26] Luiz A. de Lima, Jair M. Abe, Angel A. G. Martinez, Liliam Sayuri Sakamoto, Luigi Pavarini de Lima. “Application of architecture using AI in the training of a set of pixels of the image at aid decision-making diagnostic câncer”. 25th International Conference on Knowledge Based and Intelligent Information and Engineering Systems (KES2021). 8th – 10th September 2021 | Szczecin, Poland & Virtual. IS27: Reasoning-based Intelligent Applied Systems: HTTP://KES2021.KESINTERNATIONAL.ORG/CMSISDISPLAY.PHP

[27] De Lima, A.W.B., de Lima, L.A., Abe, J.M., Gonçalves, R.F., Alves, D. and Nakamatsu, K. “Paraconsistent Annotated Logic Artificial Intelligence Study In Support Of Manager Decision-Making”. IN: THE 2ND INTERNATIONAL CONFERENCE, 2018, BARCELONA. PROCEEDINGS OF THE 2ND INTERNATIONAL CONFERENCE ON BUSINESS AND INFORMATION MANAGEMENT - ICBIM’ 18. BARCELONA, SPAIN: ACM DL, 2018. P. 154- 157. DOI:dx.doi.org/10.1145/3278252.3278269. https://dl.acm.org/doi/10.1145/3278252.3278269

[28] Luiz A. Lima, Jair M. Abe, Angel A. G. Martinez, Jonatas S. Souza, Flávio A. Bernardini, Nilson A. Souza and Liliam S. Sakamoto. “Study of PANN Components in Image Treatment for Medical Diagnostic Decision-Making”. N.70. The 2nd International Conference on Network Enterprises & Logistics Management - NETLOG 2021. ISSN 2595-0738. HTTP://WWW.NETLOGCONFERENCE.COM/PAPERS.HTML

[29] Jonas P. Da Silva, Jair M. Abe, Luiz A. De Lima, Felipe S. David De Oliveira, Kazumi Nakamatsu. “Use of Software Metrics to Scope Control in IT Projects Using Paraconsistent Logic”. Journal WSEAS Transactions on Computer Research. WSEAS Transactions on Computer Research, ISSN/E-ISSN:1991-8755/2415-1521, Volume 6, 2018, Art. #8, pp. 55-59. (2018). https://www.wseas.org/multimedia/journals/computerresearch/2018/a145918-057.php

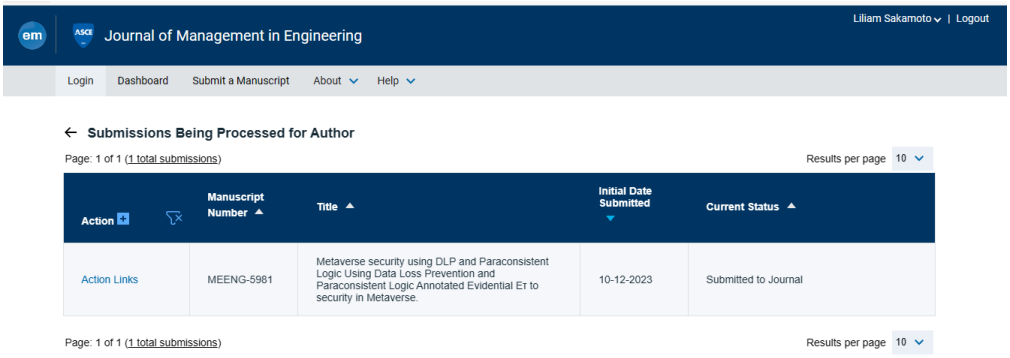
[30] Hugo Gava Insua, Jair M. Abe and Luiz A.de Lima. “Produtividade da Polícia Civil do Estado de São Paulo: uma Análise”. IJDR-International Journal of Development Research. ISSN: 2230-9926, Volume:12, Article ID:23962, 4 pages, Research Article. 2022. HTTPS://DOI.ORG/10.37118/IJDR.23962.02.2022

[31] C. Z. Kirilo, J. M. Abe, M. Nogueira, K. Nakamatsu, L. C. Machi Lozano and L. A. de lima. “Evaluation of Adherence to The Model Six Sigma Using Paraconsistent Logic”, 2018 Innovations in Intelligent Systems and Applications (INISTA), Thessaloniki, Greece, 2018, PP. 1-7, INSPEC Accession Number: 18098170, Date Added to IEEE Xplore: September 20 2018, DOI:10.1109/inista.2018.8466287. https://ieeexplore.ieee.org/document/8466287

[32] De Alencar Nääs, Irenilza; Duarte da Silva Lima, Nilsa; Franco Gonçalves, Rodrigo; Antonio de Lima, Luiz; Ungaro, Henry; Minoro Abe, Jair. “Lameness prediction in broiler chicken using a machine learning technique”. INFORMATION PROCESSING IN AGRICULTURE, v. 1, p. ‘-13, 2020. DOI: /doi.org/10.1016/j.inpa.2020.10.003 https://linkinghub.elsevier.com/retrieve/pii/S2214317320302092

ANEXO IV: Metaverse security using DLP and Paraconsistent Logic

Ano	Artigo	Status
2023	*Metaverse security using DLP and Paraconsistent Logic*	Liliam Sayuri Sakamoto, Jair Minoro Abe, Luiz Antonio de Lima.
	Submetido para revista A1	*Journal of Management in Engineering*
	SiteScore	10.4
	Highest Percentile	99% 1/57 Industrial Relations



Metaverse security using DLP and Paraconsistent Logic

Using Data Loss Prevention and Paraconsistent Logic Annotated Evidential Eτ to security in Metaverse.

First Author's Liliam, LSS, and Sakamoto*

Graduate Program in Production Engineering, Paulista University, liliam.sakamoto@gmail.com

Second Author's Jair, JMA, and Abe

Graduate Program in Production Engineering, Paulista University, airjabe@uol.com.br

Third Author's Luiz, LAL, and Lima

Graduate Program in Production Engineering, Paulista University, aula.prof.luiz@gmail.com

Objective: A research framework to optimize the security of digital assets with NFT- Non-fungible Tokens in Metaverse with the use of DLP – Data Loss Prevention and Paraconsistent Logic. Originality: Monitoring actively identifying the loss, theft, misuse, and leakage of NFT assets during their use in Metaverse, not only in a preventive way. Method: Exploratory research was carried out with the collection of an anonymized sample of a period of data from a transport company, and these data were analyzed by a program in Python in conjunction with Paraconsistent Logic. Results: Through simple monitoring, it was detected that 22% were disregarded in this process, however with the use of DLP - Data Loss Prevention in conjunction with the Paraconsistent Logic Annotated Evidential Eτ, the minimization was 22%, that is, an optimization of 11% in the security analysis.

Additional Keywords and Phrases: DLP – Data Loss Prevention, Paraconsistent Logic Annotated Evidential Eτ, Metaverse, NFT - Non-fungible Tokens.

1.Introduction

This article was developed to optimize the security of the Metaverse, aiming to minimize the risks presented by the OECD,2022. Digitally structured data can be identified in NFT - Non-fungible Tokens, as it is unique, individual, and can be stored within a database server of the environment structure. It is interesting how technological innovation with the use of the Metaverse environment, implemented the evolution in the capture of data and the transformation of them into valid information such as the use of this asset (Skalidis,2022).

However, how to ensure the security of all assets with valid NFT, without preventing their loss, theft, or leakage of these assets, and even identifying those that have invalid NFT, as well as the transit from one point to the other, in and out of the Metaverse in a lawful way (Bourlakis et al, 2009).

The application of DLP – Data Loss Prevention delimits the flow of these assets with NFT monitoring their input, if there is some transformation, until their output, where the traffic in the Metaverse environment is safe with the traceability of the audit track and the logs of the program in Python because there is active monitoring, and that alerts about loss, theft or leakage can be analyzed through Artificial Intelligence algorithms (Priest, 2018 and De Lima, 2018).

Although this type of tool is advanced in contrast to simple daily manual monitoring, some alerts can present doubts and be disjointed, causing problems in traceability. For this reason, another layer of analysis was implemented with the Paraconsistent Logic Evidential Annotated Et adds greater assertiveness in decision-making (Insua, Abe & Lima, 2022). An exploratory survey was carried out with the collection of a sample of a transport company that asks for confidentiality, since the data were anonymized in accordance with the LGPD – General Data Protection Law of Brazil to test this implementation, referring only to a period of data (De Souza et al, 2019).

The differential of each organization in the monitoring and identification of these assets is in the ability to manage risk in an automated way with artificial intelligence technologies, or as in this proposal with a DLP implementation because the more the Metaverse is accessed and its information transformed with or without user interaction, an asset can present an increasing value, therefore, the greater the importance in the security aspect (Silowash, 2013).

1.1 Metaverse

The curiosity of humanity in search of "parallel worlds" has always been a challenge, and with the creation of the Metaverse comes an opportunity to experience and live in a cyberspace or digital environment. This type of experience leads to a Networked Society experience (Schlemmer, 2008), such a stimulus arouses curiosity in the individuals of the population.

The Metaverse as an emerging technology is reshaping society through differentiated opportunities, but it also presents risks, such as violation of privacy, a threat to security, and virtual theft, with the need to protect established rights (OECD, 2022).

This materialization of parallel worlds on the internet, the Metaverse constitutes the vision of the progress of technology to create an immersive social and organizational life with the use of artificial intelligence (Samarnggoon, 2023).

Several Metaverse models have been developed as practical platforms based on youth games (La Fuente Prieto, 2022), such as Second Life, Meta Horizons (Facebook), Fortnite, Pokemon GO, Spatial, Sandbox, Somnium Space, Decentraland, Cryptovoxels, Axie Infinity, Alien Worlds, Illuvium, Bloktopia, Star Atlas, Roblox. For example, Somnium Space features unique wallets that buy or sell NFTs (Ante, 2022).

The Metaverse environment with the use of visual stimuli causes an increase in emotional states that reaches an experience of almost reality, proven by the analysis of stimuli even with the support of Electroencephalography, as presented by Daşdemir, 2022.

Some Museums already present the exhibition content in Metaverse, combining augmented reality and a virtual world (Choi, 2017). Other companies are using it for events to promote their brands, such as Gucci, which within games transacts the sale of products (virtual assets with NFT) like your bags even with more expensive value than the real product in your stores (Forbes,2022).

There is also the approach that in the future everyone's personal data will have an NFT that would identify it as unique to that person and non-transferable, which would prevent the undue commercialization of that information, especially if it contains sensitive data such as medical information, making it possible to also identify its veracity and due use (Skalidis, 2022).

This is an example of a training meeting room for LGPD in Metaverse with the participation of DPOs from the APDADOS Scientific Committee, using the Spatial application:

Figure 1: Spatial Metaverse – Scientific Committee, via Authors.

In most studies referring to the Metaverse, it is observed that there is a need to develop policies that deal with the environment in a regulatory manner and even only in terms of information security (OCDE, 2022 and Bourlakis et al, 2009).

1.2 DLP - Data Loss Prevention

The concept of DLP – Data Loss Prevention is not new, but it is being used more nowadays, as ready-made tools on the market already have some functionalities of this built-in monitoring standard. Applications such as Microsoft Office 365 are already capable of bringing some internally standardized templates for compliance with the LGPD or other regulations such as GDPR – General Data Protection Regulation (European Commission, 2016). Even trying to carry out an analysis for compliance with the standards, with the lack of depth in the evaluation of safety criteria, data loss may occur (Silowash, 2013).

Trend Micro's antivirus features a solution called Apex One, which is a SaaS – Systems as a Service for endpoint and server security, with the functionality to report logs on DLP – Data Loss Prevention, which can be configured according to the needs of the company, for example, search for disclosure of information on CPF, CNPJ, Invoices, credit card, social security, NFT or even specific words, etc. (Silowash, 2013).

Each of these tools uses artificial intelligence algorithms that cross data and monitor patterns pre-formatted by the information security analyst, starting from classical logic for conclusions and presentations for decision-making. Mainly, when a situation of data loss, theft, or leakage is detected.

The DLP must have an assertive configuration to capture results aligned with the company's needs, such as the number of credit card transactions per customer, which ones are valid or if there are blockages for the same card, among other situations, to evaluate suspicious situations (Priest, 2018).

DLP can also be standardized to evaluate suspected malware, ransomware, and spam, using public safety lists such as the Mitre Attack website (Sikorski, 2012).

Many companies have developed tools that manage to unify alerts about malicious activities such as IP scanning, data flow, malicious emails, intrusion attempts, firewall monitoring, review of security policy updates, and implement security patches, but in this study, the focus is specifically on loss, theft, or leakage of NFT-enabled assets within the Metaverse environment (Priest, 2018).

In addition, optimizing the effective identification of redundant, invalid, and non-NFT assets, and their identification, if any change or undue transition has occurred, benefits the effective monitoring of these valuable assets (Silowash, 2013). DLP tools require testing and calibration, a situation in which alerts with excessive false positives can occur, but they use classical logic (Akama, 2016).

The level of sophistication reached by malicious software requires constant efforts to mitigate this practice with preventive solutions, such as the use of DLP for effective monitoring in conjunction with Paraconsistent Logic, to innovate and be one step ahead of them (Singh,2018).

1.3Paraconsistent Logic Annotated Evidential Eτ

Annotated logics constitute a class of paraconsistent logic. Such logics are related to certain complete lattices, which play an essential role. A knowledge of an expert on an analyzed subject, questions are used to capture opinions that are normalized in logic between 0 and 1, as shown in Figure 5. These values are respectively the favorable evidence that is expressed by the symbol μ and the contrary evidence by λ. Logic Eτ must follow the process (look at figure 1) during the application, which can be seen in the figure:

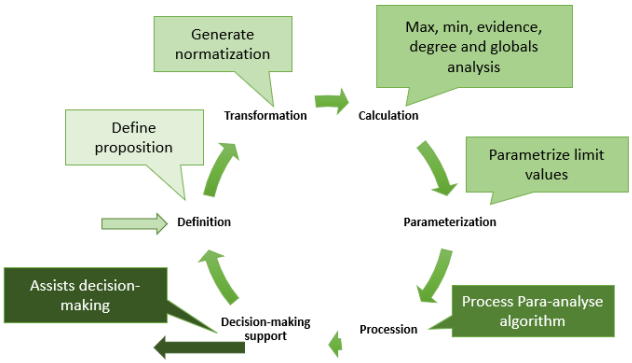


Figure 1: Steps Process Paraconsistent Logic Annotated Evidential Eτ, via Authors.

The definition means understanding and creating the proposition to be understood and which should reflect the problem. The transformation would be treating data from favorable and unfavorable evidence. We usually normalize the answer between the number zero and one to be handled by Logic Eτ. Thus, it becomes possible to perform the processing of calculations of the data collected. With this, we have the Favorable Degree and the Degree of Unfavorable. The acceptable limits are parameterized to obtain an analysis that makes the data and can be transformed into useful information (De Alencar Nääs, 2020). The intelligence is given when applying the para-analyzer algorithm that contains all the information to execute the Eτ Logic.

As summarized by Abe et al., 2011, programs can now be built using paraconsistent logic, making it possible to treat inconsistencies directly and elegantly. This feature can be applied in specialist systems, object-oriented databases, representation of contradictory knowledge, etc., with all the implications in artificial intelligence.

In Abe et al., 2011: "Paraconsistent Logic Annotated Evidential Eτ has an Eτ language, and the atomic propositions are of the type p (μ, λ) where p is a proposition and μ, λ∈ [0, 1]. Intuitively, μ indicates the degree of unfavorable evidence1 of p and λ the degree of contrary evidence of p. The reading of the values μ, λ depends on the applications considered and may change in fact μ it may be the degree of belief favorable and λ it may be the degree of belief contrary to proposition p; also, μ can indicate the probability expressed by p occurring and λ the improbability expressed by p occurring. The atomic propositions p (μ, λ) of logic Eτ can intuitively be read as: I believe in p with the degree of favorable belief μ and the degree of contrary belief λ, or the degree of favorable evidence of p is μ and the degree of evidence to the contrary of p is λ". Paraconsistent logics can serve as underlying logic of theories in which A and ¬A (the negation of A) are both true without being trivial (Abe & Nakamatsu, 2015). There are many types of

paraconsistent systems. In this text, it considers the Paraconsistent Logic Annotated Evidential Eτ (Dill, 2014). The formulation in Logic Eτ is of the type p (μ, λ), in which p is a proposition and e (μ, λ) ϵ [0, 1] is the real unitary closed interval.

A proposition p (μ, λ) can be read as: " The favorable evidence of p is μ and the unfavorable evidence is λ" (Abe, 2011). For instance, p (1.0, 0.0) can be read as a true proposition, p (0.0,1.0) as false, p (1.0, 1.0) as inconsistent, p (0.0, 0.0) as Paracomplete, and p (0.5, 0.5) as an indefinite proposition (Abe et al, 2011). Also, to introduce the following concepts: Uncertainty degree: Gun (μ, λ) = μ + λ - 1 ($0 \leq \mu, \lambda \leq 1$) and Certainty degree: Gce (μ, λ) = μ - λ ($0 \leq \mu, \lambda \leq 1$) [9]. An order relation is defined on [0, 1]: ($\mu 1, \lambda 1$) ≤ ($\mu 2, \lambda 2$) ↔ $\mu 1 \leq \mu 2$ and $\lambda 2 \leq \lambda 1$, constituting a lattice that will be symbolized by τ.

With the uncertainty and certainty degrees, we can get the following 12 output states (Table 2): extreme and non-extreme states. It is worth observed that this division can be modified according to each application (SUBRAHMANIAN, 1987). Para-analyzer Algorithm, this proposed algorithm, there is a set of information obtained, which can sometimes seem contradictory, making it difficult to analyze the scenario for risk analysis. Generally, in such situations, this information is discarded or ignored, that is, they are considered "dirty" of the system, however at best they may even receive different treatment. Silva Filho, Abe, and Torres, 2011 quote: "However, the contradiction most of the time contains decisive information, as it is like the encounter of two strands of opposing truth values. Therefore, to neglect it is to proceed in an anachronistic way, and that is why we must look for languages that can live with the contradiction without disturbing the other information. As for uncertainty, we must think of a language that can capture the 'maximum' of 'information' of the concept". In this line of reasoning for the analysis based on Paraconsistent Logic, situations of Inconsistency and Para completeness will be considered together with the True and False, represented according to Table 1:

Table 1: Extreme States - Abe et al. (2011)

Extreme States	Symbol
True	V
False	F
Inconsistent	T
Paracomplete	⊥

The set of these states or objects (τ = {F, V, T,⊥}) can also be called annotation constants and can be represented using the Hasse diagram as shown in figure 2 (Kirilo, 2018):

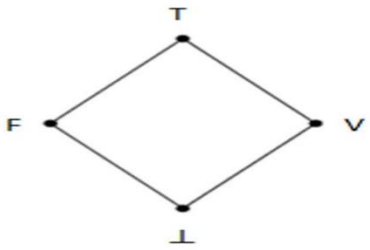


Figure 2: Hasse diagram, via Authors.

"The operator about τ é: ~:|τ| →|τ| that will operate, intuitively, like this:

~ T = T (the 'negation' of an inconsistent proposition is inconsistent)

~ V = F (the 'negation' of a true proposition is false)

~ F = V (the 'negation' of a false proposition is true)

~⊥ = ⊥ (the 'negation' of a Paracomplete proposition is Paracomplete)

Paraconsistent Logic Annotated Evidential Eτ will be used; this type must be composed of 1, 2 or "n" values.

With the calculations of the values of the axes that make up the representative figure of the lattice, it can be divided or internally delimited into several regions of different sizes and formats, thus obtaining a discretization of the same.

From the bounded regions of the lattice, it is possible to relate the resulting logical states, which, in turn, will be obtained by interpolating the Degrees of Certainty Gc and Contradiction Gct. Thus, for each interpolation between the degrees of certainty and contradiction, it is possible to extract information to assist in decision making (SUBRAHMANIAN, 1987).

The representation of table 2 shows a representation of the lattice constructed with values of Degrees of Certainty and Contradiction and sectioned into 12 states. Thus, at the end of the analysis, one of the 12 possible resulting logical states will be obtained as an answer for decision making (Carvalho, 2011 and Abe,2011).

Table 2. Non-extreme states - Abe et al. (2011)

Non-extreme States	Symbol
Quasi-true tending to Inconsistent	$QV \rightarrow T$
Quasi-true tending to Paracomplete	$QV \rightarrow \perp$
Quasi-false tending to Inconsistent	$QF \rightarrow T$
Quasi-false tending to Paracomplete	$QF \rightarrow \perp$
Quasi-inconsistent tending to True	$QT \rightarrow V$
Quasi-inconsistent tending to False	$QT \rightarrow F$
Quasi-Paracomplete tending to True	$Q\perp \rightarrow V$
Quasi-Paracomplete tending to false	$Q\perp \rightarrow F$

Some additional control values are:

- Vscct = maximum value of uncertainty control = Ftun
- Vscc = maximum value of certainty control = Ftce
- Vicct = minimum value of uncertainty control = -Ftun
- Vicc = minimum value of certainty control = -Ftce

All states are represented in the next figure 5.

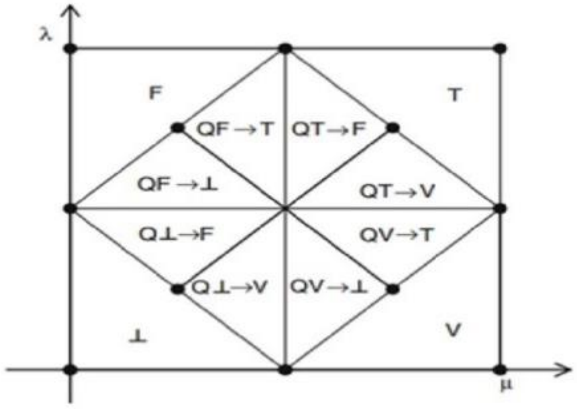


Figure 5: Non-extreme states, via Abe et al. (2011).

2. OPTIMIZATION OF METAVERSE SECURITY

When using the concept of DLP – Data Loss Prevention (see Figure 6), the following granularity was considered to obtain graduation for the types of data loss detected (Bioni, 2019):

2.1DLP with Paraconsistent Logic Structure

DLP – Data Loss Prevention helps to delimit:

• Level 1 – NFT assets.

• Level 2 – location of assets within the Metaverse environment.

• Level 3 – detection of invalid or incomplete assets.

• Level 4 – application of Paraconsistent Logic Et.

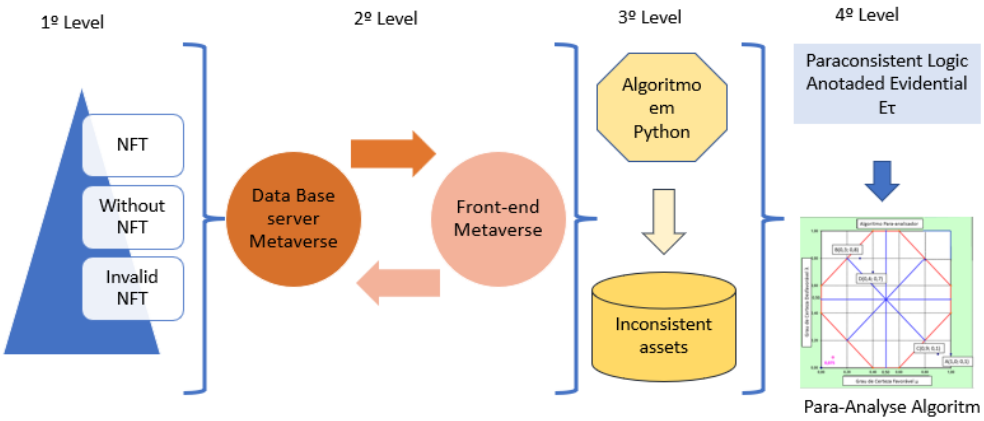


Figure 6: DLP – Data Loss Prevention using Eτ Evidential Annotated Paraconsistent Logic, via Authors.

- Level 1 – NFT assets:
  - what is your origin?
  - where they are stored, if they are in the structure's Database.
  - does it go through NFT consistency and validation?
  - How is the asset exit transition made with NFT?
  - Detection of assets with invalid NFT.
  - Audit trail and transaction log.
- Level 2 – location of assets within the Metaverse environment:
  - on the Network (on the database server or application server).
  - On the front end (on the web server).
- Level 3 – detection of invalid or incomplete assets.
  - Validation of the mass of data collected from the carrier company for just one period to be analyzed by the Python Algorithm to verify valid and contradictory assets.
- Level 4 – application of Paraconsistent Logic Eτ.
  - The exception information presented by the Python algorithm report is aligned. They are inserted into the analyzer of the para-analyzer Algorithm to identify the greatest assertiveness in decision-making.

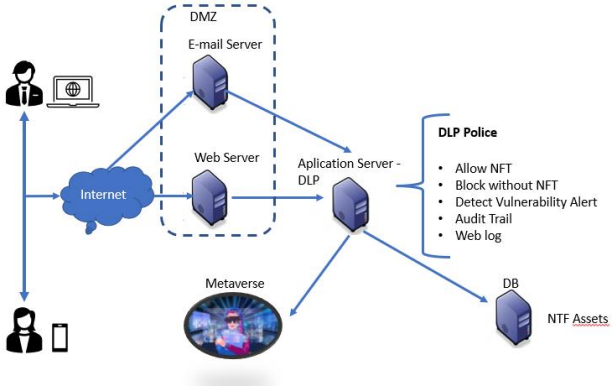


Figure 7: DLP Structure in Metaverse, via Authors.

2.2 Use of Python Program

The mass of data to be tested had the following layout (figure 8) with this caption, with python program:

The First Level (asset type):

- Assets with NFT = 1.1
- Assets without NFT =1.2
- Assets with invalid NFT =1.3

The second level (location):

- Within the Metaverse Environment Database = 2.1
- Inside the Metaverse frontend = 2.2

The third level (validation):

- valid assets =3.1
- contradictory assets=3.2

The fourth level (paraconsistent logic):

- True=4.1
- False=4.2
- Incomplete=4.3
- Paracomplete=4.4

For which each (see figure 8) of the captured lines presented a description of the source of the assets (Li, 2020).

Level 1	Level 2	Level 3	Level 4
1.1	2.1	3.1	4.1.
1.2	2.2	3.2	4.2
1.3	2.1	3.2	4.4
1.1	2.1	3.2	4.1
1.2	2.1	3.2	4.3
1.3	2.1	3.2	4.2
1.1	2.1	3.2	4.1
1.2	2.1	3.1	4.3
1.3	2.1	3.1	4.2
1.1	2.1	3.1	4.1
1.2	2.1	3.1	4.3
1.3	2.1	3.1	4.2
1.1	2.1	3.1	4.1
1.2	2.1	3.1	4.3
1.3	2.1	3.1	4.4
1.1	2.2	3.1	4.1
1.2	2.2	3.1	4.3
1.3	2.2	3.1	4.4
1.1	2.2	3.1	4.1
1.2	2.2	3.1	4.3

Figure 8: Example of mass data, via Authors.

2.3 Python Program

The OECD document (2023) addresses that AI enables and supports emerging technologies, such as immersive virtual environments, such as “virtual worlds”, such as the “metaverse”. In this context, an adaptation of algorithms in Python (Kopec, 2019) was developed and validated in practice with supervised learning to meet the necessary parameterizations (Lutz, 2013 and McKinney, 2022). This computational algorithm in Python used in level 3:

Python Program: NFT

```
import csv
from typing import List
from util import normalize_by_feature_scaling
from network import Network
from random import shuffle

if __name__ == "__main__":
    nft_parameters: List[List[float]] = []
    nft_classifications: List[List[float]] = []
    nft_species: List[int] = []
    with open('C:/Users/luizl/Downloads/RedeNeural_01/nft.csv', mode='r') as nft_file
        nfts: List = list(csv.reader(nft_file, quoting=csv.QUOTE_NONNUMERIC))
        shuffle(nfts) # get our lines of data in random order
        for nft in nfts:
            parameters: List[float] = [float(n) for n in nft[1:14]]
            nft_parameters.append(parameters)
            species: int = int(nft[0])
            if species == 1:
                nft_classifications.append([1.0, 0.0, 0.0])
            elif species == 2:
                nft_classifications.append([0.0, 1.0, 0.0])
            else:
                nft_classifications.append([0.0, 0.0, 1.0])
            nft_species.append(species)
    normalize_by_feature_scaling(nft_parameters)

# Lista [camada de entrada,camada intermediaria, camada de saida], taxa de Aprendizag
    nft_network: Network = Network([13, 7, 3], 0.9)
    def nft_interpret_output(output: List[float]) -> int:
        if max(output) == output[0]:
```

2.4 Paraconsistent Logic Annotated Evidential Eτ

The exception information presented by the Python algorithm report represents about 22% and should be discarded from the analyses, however, it is inserted into the Para-Analyzer Algorithm to identify to create a second layer of verification based on artificial intelligence for greater assertiveness in decision-making.

The Paraconsistent Logic Annotated Evidential Eτ considers a proposition being represented by annotation values. According to this concept, an algorithm called a para-analyzer was created (ABE, 2011).

The defined proposition was “Is regulation with the use of DLP together with Paraconsistent Logic a security implementation for the Metaverse?”

Eighteen specialists in the areas of Information Technology and Data Protection are selected as knowledge engineers so that their analyzes on the defined items are scored within the para-analyzer algorithm referring to the proposition.

IT and Data Protection professionals were selected as Knowledge Engineers - CE, and each of them received a form to answer the: Degree of favorable evidence μ and Degree of unfavorable evidence λ for each of the factors, within the process of transport specifically for the delimitation of routes (Akama, 2016 and Carvalho, 2011).

Table 3: Factors and Sectors

Order	Weight	Factors	Sector
A	1	F1: Are the assets standardized with valid NFTs in the metaverse environment?	S1: Inside Data Base
B	1	F1: Are the assets standardized with valid NFTs in the metaverse environment?	S2: Inside Front-end
C	2	F2: Are there assets without NFTs in the metaverse environment?	S1: Inside Data Base

D	2	F2: Are there assets without NFTs in the metaverse environment?	S2: Inside Front-end

Table 4: Knowledge engineers' consideration of F1-S1 and F1-S2

KE	Factor	Sector	Weight	Favourable Degree of Evidence μ	Unfavourable Degree of Evidence λ
1(Group1)	F1	S1	1	1.0	0.1
2(Group1)	F1	S1	1	0.9	0.2
3(Group1)	F1	S1	1	0.9	0.0
4(Group1)	F1	S1	1	0.9	0.1
5(Group1)	F1	S1	1	0.9	0.2
6(Group1)	F1	S1	1	0.8	0.3
7(Group1)	F1	S1	1	0.9	0.0
8(Group1)	F1	S1	1	0.9	0.1
9(Group1)	F1	S1	1	0.9	0.0
10(Group2)	F1	S2	1	0.4	0.2
11(Group2)	F1	S2	1	0.5	0.1
12(Group2)	F1	S2	1	0.3	0.2
13(Group2)	F1	S2	1	0.7	0.6
14(Group2)	F1	S2	1	0.6	0.1
15(Group2)	F1	S2	1	0.5	0.2
16(Group2)	F1	S2	1	0.3	0.3
17(Group2)	F1	S2	1	0.4	0.0
18(Group2)	F1	S2	1	0.4	0.1

Table 5: Knowledge engineers' consideration of F2-S1 and F2-S2

KE	Factor	Sector	Weight	Favourable Degree of Evidence μ	Unfavourable Degree of Evidence λ
1(Group1)	F2	S1	2	0.6	0.5
2(Group1)	F2	S1	2	0.8	0.7
3(Group1)	F2	S1	2	0.7	0.2
4(Group1)	F2	S1	2	0.8	0.3
5(Group1)	F2	S1	2	0.8	0.1
6(Group1)	F2	S1	2	0.9	0.2
7(Group1)	F2	S1	2	0.8	0.1
8(Group1)	F2	S1	2	0.7	0.2
9(Group1)	F2	S1	2	1.0	0.1
10(Group2)	F2	S2	2	0.1	1.0
11(Group2)	F2	S2	2	0.3	0.9
12(Group2)	F2	S2	2	0.1	0.8
13(Group2)	F2	S2	2	0.1	0.7
14(Group2)	F2	S2	2	0.2	0.9
15(Group2)	F2	S2	2	0.1	1.0
16(Group2)	F2	S2	2	0.3	0.8
17(Group2)	F2	S2	2	0.1	0.9
18(Group2)	F2	S2	2	0.0	0.9

Table 7: Differential after using Paraconsistent Logic of the mass of discarded data

Result	Factor	Section	Conclusion		Definition	Level 4
			μ	λ		
A	F1	S1	0,9	0,1	True	4.1
B	F1	S2	0,4	0,8	Incomplete	4.3
C	F2	S1	0,8	0,2	True	4.1
D	F2	S2	0,2	0,8	False	4.2

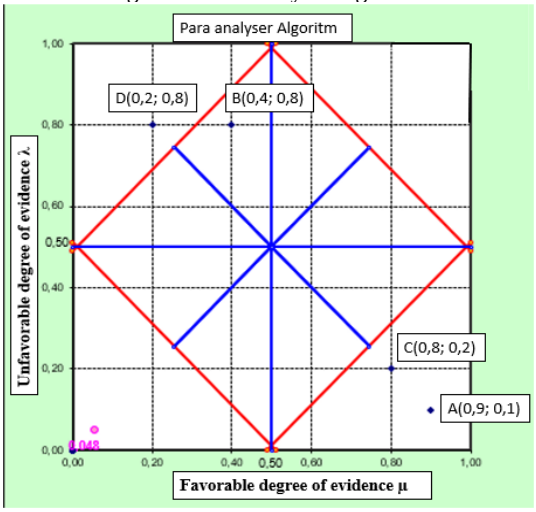
3. Results

The mass of captured data presented about 1 to 2 pieces of information per day for a month, comprising an amount for analysis of 28 pieces of data, of which the algorithm verified 78% in the Python Program as effective and 22% as contradictory. Data considered contradictory were passed on to the Para-analyzer algorithm (figure 9), obtaining a differential of 5% true, 6% false, 5% incomplete, and 6% para-complete.

Fist test:

However, the interval considered false may undergo a more demanding screening that requires more effective monitoring due to the percentage of occurrences that could be an invisible data leak alert, but which occur subtly.

Figure 9: Para-analyzer algorithm



The analyzed assets were discarded after going through the Python program. However, they were reanalyzed by the para-analyzer algorithm, which optimized over 22% of the assets (5% true and 6% false), minimizing the loss of this analysis by only 11%:

Table 7: Differential after using Paraconsistent Logic

Percentual of reach	State	Conclusion
5%	True	1 assets
6%	False	2 assets

5%	Incomplete	1 assets
6%	Paracomplete	2 assets

4. Conclusion

It is concluded that the alignment between the use of DLP - Data Loss Prevention with the Paraconsistent Logic presents a significant gain for the issue of monitoring and analysis of the loss of NFT assets in the Metaverse environment.

The implementation of this type of tool can increase Metaverse's performance in relation to the level of information security, which used to despise 22% of the data as inconsequential to only 11% (5% true and 6% false), considering that even these intervals can be monitored so that their effectiveness of 100% of the data transited, kept stored and sent by email or analyzed social networks.

It is understood that the combination of data privacy professionals with the improvement of tools such as DLP with the use of Paraconsistent Logic brings the possibility of significant advances in the adequacy of the LGPD and an increase in the flow of new businesses within the environment of the Metaverse safely.

ACKNOWLEDGMENTS

We thank the research group Paraconsistent logic and artificial intelligence maintained by Paulista University and conducted by researcher Dr. Abe. This study was financed in part by the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) -Financial Code 001.

REFERENCES

Abe, J.M., et al.: Lógica Paraconsistente Anotada Evidencial Et, pp. 38–39. Comunnicar, Santos, 2011.

Abe, J. M. Nakamatsu K. Introduction to Annotated Logics - Foundations for Paracomplete and Paraconsistent Reasoning, Series Title Intelligent Systems Reference Library, Volume 88, Publisher Springer International Publishing, Copyright Holder Springer International Publishing Switzerland, eBook ISBN 978-3-319-17912-4, DOI 10.1007/978-3-319-17912-4, Hardcover ISBN 978-3-319-17911-7, Series ISSN 1868-4394, Edition Number 1, 190 pages, 2015.

Akama, S. Towards Paraconsistent Engineering, Intelligent Systems Reference Library, Germany: Springer (2016).

AI, T. (2023). OECD DIGITAL ECONOMY PAPERS. Harnessing the power of AI and emerging Technologies: Background paper for the CDEP Ministerial meeting (https://www.oecd-ilibrary.org/docserver/f94df8ec-en.pdf?expires=1685828058&id=id&accname=guest&checksum=7BDAF5EA7F8459F867B95BA5118521E8)

Ante, L. (2022). Non-fungible token (NFT) markets on the Ethereum blockchain: Temporal development, cointegration and interrelations. Economics of Innovation and New Technology, 1-19.

Bioni, Bruno Ricardo. Proteção de Dados Pessoais: a função e os limites do consentimento. Ed 1. Vol. único. Rio de Janeiro: Foresee, 2019

Bourlakis, M., Papagiannidis, S., & Li, F. (2009). Retail spatial evolution: paving the way from traditional to metaverse retailing. Electronic Commerce Research, 9, 135-148.

Brazil. Lei Geral de Proteção de Dados Pessoais (LGPD). Lei nº 13.709, de 14 de agosto de 2018. Available in :http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm . Accessed on: 21/04/2023.

Carvalho, Fábio Romeu de; ABE, Jair Minoro. Tomadas de Decisão com Ferramentas da Lógica Paraconsistente Anotada: Método Paraconsistente de Decisão - MPD. 1a edição. São Paulo: Blucher, 2011

Choi, H. S., & Kim, S. H. (2017). A content service deployment plan for metaverse museum exhibitions—Centering on the combination of beacons and HMDs. International Journal of Information Management, 37(1), 1519-1527.

Daşdemir, Y. (2022). Cognitive investigation on the effect of augmented reality-based reading on emotion classification performance: A new dataset. Biomedical Signal Processing and Control, 78, 103942.

Da Silva, Jonas P. et al. "Use of Software Metrics to Scope Control in IT Projects Using Paraconsistent Logic". Journal WSEAS Transactions on Computer Research. WSEAS Transactions on Computer Research, ISSN/E-ISSN:1991-8755/2415-1521, Volume 6, 2018, Art. #8, pp. 55-59. (2018). https://www.wseas.org/multimedia/journals/computerresearch/2018/a145918-057.php

De Alencar Nääs, Irenilza; Duarte da Silva Lima, Nilsa; Franco Gonçalves, Rodrigo; Antonio de Lima, Luiz; Ungaro, Henry; Minoro Abe, Jair. "Lameness prediction in broiler chicken using a machine learning technique". INFORMATION PROCESSING IN AGRICULTURE, v. 1, p. '-13, 2020. DOI: /doi.org/10.1016/j.inpa.2020.10.003 https://linkinghub.elsevier.com/retrieve/pii/S2214317320302092

De Carvalho, F.R., Abe, J.M.: Tomadas de decisão com ferramentas da Lógica Paraconsistente Anotada. São Paulo. Blucher, pp. 37–47, 2011.

De Carvalho, F.R., Brunstein, I., Abe, J. M.: Paraconsistent Annotated Logic in Analysis of Viability: in approach to product launching. In: Dubois, D.M. (ed.), vol. 718, pp. 282–291, 2011.

Dill, R.P., Da Costa Jr., N., Santos, A. A. P.: Corporate Profitability Analysis: A Novel Application for Paraconsistent Logic. Applied Mathematical Sciences 8, 2014.

De La Fuente Prieto, J., Lacasa, P., & Martínez-Borda, R. (2022). Approaching metaverses: Mixed reality interfaces in youth media platforms. New Techno Humanities.

De Lima L.A., Abe J.M., Martinez A.A.G., de Frederico A.C., Nakamatsu K., Santos J. "Process and Subprocess Studies to Implement the Paraconsistent Artificial Neural Networks for Decision-Making". In: Jain V., Patnaik S., Popențiu Vlădicescu F., Sethi I. (Eds) Recent Trends in Intelligent Computing, Communication and Devices. Advances in Intelligent Systems nd Computing, Vol 1006. Springer, Singapore. 2019 Print ISBN: 978-981- 13-9405-8; Online Isbn: 978-981-13-9406-5; HTTPS://DOI.ORG/10.1007/978-981-13-9406-5_61

De Lima, Luiz A.; Abe, Jair M.; Kirilo, Caique Z.; Da Silva, Jonas P.; Nakamatsu, Kazumi. "Using Logic Concepts in Software Measurement." PROCEDIA COMPUTER SCIENCE, v. 131, p. 600-607, 2018. http://dx.doi.org/10.1016/j.procs.2018.04.302

De Lima, A.W.B., de Lima, L.A., Abe, J.M., Gonçalves, R.F., Alves, D. and Nakamatsu, K. "Paraconsistent Annotated Logic Artificial Intelligence Study In Support Of Manager Decision-Making". IN: THE 2ND INTERNATIONAL CONFERENCE, 2018, BARCELONA. PROCEEDINGS

OF THE 2ND INTERNATIONAL CONFERENCE ON BUSINESS AND INFORMATION MANAGEMENT - ICBIM' 18. BARCELONA, SPAIN: ACM DL, 2018. P. 154- 157. DOI:dx.doi.org/10.1145/3278252.3278269. https://dl.acm.org/doi/10.1145/3278252.3278269

Dowling, M. (2022). Is non-fungible token pricing driven by cryptocurrencies?. Finance Research Letters, 44, 102097.

Dowling2, M. (2022). Fertile LAND: Pricing non-fungible tokens. Finance Research Letters, 44, 10296.

European Comission GDPR - General Data Protection Regulation 2016 Available in :< https://gdpr-info.eu/ > Accessed on: 21/04/2023.

Forbes. Available in: https://forbes.com.br/forbes-tech/2022/01/exemplos-do-metaverso-marcas-que-atuam-com-propriedade/. Accessed on: 22/05/2023.

García, R., Cediel, A., Teixidó, M., & Gil, R. (2022). Semantics and non-fungible tokens for copyright management on the metaverse and beyond. arXiv preprint arXiv:2208.14174

Goanta, C. (2020). Selling LAND in Decentraland: The regime of non-fungible tokens on the Ethereum blockchain under the digital content directive. Disruptive technology, legal innovation, and the future of real estate, 139-154.

Insua, Hugo Gava, Abe, J.M. & Lima, L.A. "Produtividade da Polícia Civil do Estado de São Paulo: uma Análise". IJDR-International Journal of Development Research. ISSN: 2230-9926, Volume:12, Article ID:23962, 4 pages, Research Article. 2022. HTTPS://DOI.ORG/10.37118/IJDR.23962.02.2022

Kirilo,C. Z. et al. "Evaluation of Adherence to The Model Six Sigma Using Paraconsistent Logic", 2018 Innovations in Intelligent Systems and Applications (INISTA), Thessaloniki, Greece, 2018, PP. 1-7, INSPEC Accession Number: 18098170, Date Added to IEEE Xplore: September 20 2018, DOI:10.1109/inista.2018.8466287. https://ieeexplore.ieee.org/document/8466287

Kopec, D. (2019). Classic computer science problems in Python. Simon and Schuster.

Lamy, Eduardo de Avelar. Princípio da Fungibilidade no Processo Civil. São Paulo (SP): Dialética, 2007

Li, T., Butrovich, M., Ngom, A., Lim, W. S., McKinney, W., & Pavlo, A. (2020). Mainlining databases: Supporting fast transactional workloads on universal columnar data file formats. arXiv preprint arXiv:2004.14471.

Lima, Luiz Antônio de, et al. "DPO no Brasil sob a ótica da LGPD - Lei Geral de Proteção de Dados". Instituto EXIN - Ministry of Economic Affairs in the Netherlands. 2020. https://www.exin.com/br-pt/dpo-no-brasil-sob-a-otica-da-lgpd-lei-de-protecao-de-dados/

Lima, L. A., Abe, J. M., Martinez, a. A. G., Santos, J., Albertini, G., & Nakamatsu, K. (2019). "The Productivity Gains Achieved in Applicability of The Prototype AITOD with Paraconsistent Logic in Support in Decision-Making in Project Remeasurement". Proceedings of the 9th International Conference of Information and Communication Technology [ICICT-2019] Nanning, Guangxi, China January 11-13, 2019 (http://aivr.org/index.html). Edited by Srikanta Patnaik Volume 154, Pages 1-844 (2019). Procedia Computer Science, pp. 347–353. https://doi.org/10.1016/j.procs.2019.06.050

Lima, Luiz A. de, et al "Application of architecture using AI in the training of a set of pixels of the image at aid decision-making diagnostic câncer". 25th International Conference on Knowledge Based and Intelligent Information and Engineering Systems (KES2021). 8th – 10th September 2021 | Szczecin, Poland & Virtual. IS27: Reasoning-based Intelligent Applied Systems: HTTP://KES2021.KESINTERNATIONAL.ORG/CMSISDISPLAY.PHP

Lima, Luiz A. et al. "Study of PANN Components in Image Treatment for Medical Diagnostic Decision-Making". N.70. The 2nd International Conference on Network Enterprises & Logistics Management - NETLOG 2021. ISSN 2595-0738. HTTP://WWW.NETLOGCONFERENCE.COM/PAPERS.HTML

Lutz, M. (2013). Learning python: Powerful object-oriented programming. " O'Reilly Media, Inc.".

McKinney, W. (2022). Python for Data Analysis. " O'Reilly Media, Inc."

Priest, Graham, Koji Tanaka, and Zach Weber, "Paraconsistent Logic", The Stanford Encyclopedia of Philosophy (Summer 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2018/entries/logic-paraconsistent/>. ISSN: 1095-5054

Peixoto, F. H. (2020). Projeto Victor: relato do desenvolvimento da inteligência artificial na repercussão geral do Supremo Tribunal Federal. Revista Brasileira de Inteligência Artificial e Direito-RBIAD, 1(1), 1-22.

OECD - Organization for Economic Co-operation and Development – Harnessing the power of AI and emerging technologies - OECD Digital Economy Papers - November 2022 nº 340

Samarnggoon K, Grudpan S, Wongta N and Klaynak K. (2023). Developing a Virtual World for an Open-House Event: A Metaverse Approach.Future Internet. 10.3390/fi15040124. 15:4. (124).

Schlemmer, E., & Backes, L. (2008). Metaversos: novos espaços para construção do conhecimento. Revista Diálogo Educacional, 8(24), 519-532.

Skalidis, I., Muller, O., & Fournier, S. (2022). The metaverse in cardiovascular medicine: applications, challenges, and the role of non- fungible tokens. Can J Cardiol, 38(9), 1467-8.

Sikorski, M., Honig, A. "Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software", 2012, No Starch Press.

SILOWASH, George J.; KING, Christopher. Insider threat control: Understanding data loss prevention (DLP) and detection by correlating events from multiple sources. CARNEGIE-MELLON UNIV PITTSBURGH PA SOFTWARE ENGINEERING INST, 2013.

Singh, J., Singh, J., "Challenges of Malware Analysis: Obfuscation Techniques", 2018. Disponível em: http://50.87.218.19/ijiss/index.php/ijiss/article/view/327

Siqueira, D. P., Lara, F. C. P., & Lima, H. F. C. (2020). Acesso à justiça e inteligência artificial: abordagem a partir da revisão sistemática da literatura. Revista Argumentum-Argumentum Journal of Law, 21(3), 1265-1277.

Souza, Jonatas S. de, et al. "The General Law Principles for Protection the Personal Data and their Importance". In: 7th International Conference on Computer Science Engineering and Information Technology (CSEIT 2020 - https://arxiv.org/abs/2009.14313), 2020. Computer science & information technology (cs & it), Copenhagen, Denmark. Anais 2020. V. 10. P. 109. HTTP://DX.DOI.ORG/10.5121/CSIT.2020.101110

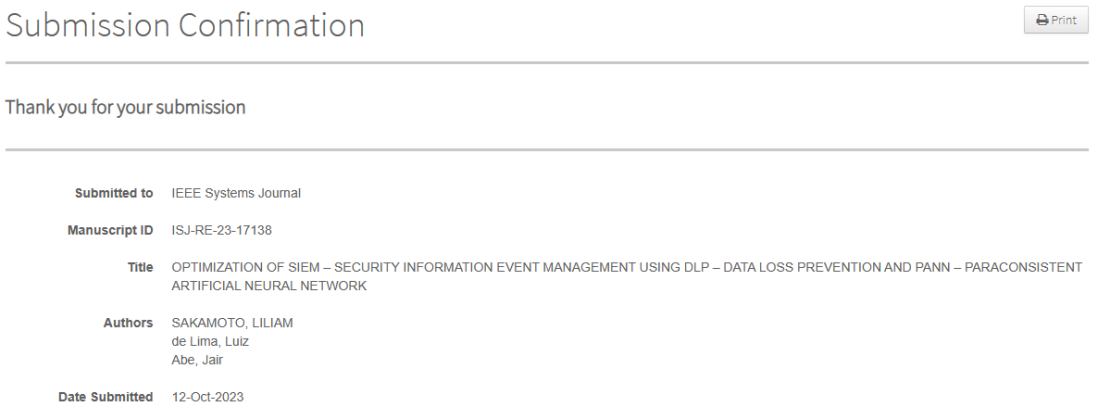
SOUZA, JONATAS S. DE, et al "The Brazilian Law on Personal Data Protection", International Journal of Network Security & Its Applications (IJNSA - https://airccse.org/journal/jnsa20_current.html), November 2019, Volume 12, Number 6, ISSN: 0974-9330[online]; 0975-2307 [Print]. https://aircconline.com/ijnsa/V12N6/12620ijnsa02.pdf

SUBRAHMANIAN, V.: On the semantics of quantitative logic programs. In: Proceedings of the 4th IEEE Symposium on Logic Programming, pp. 173–182 (1987)

Wang, Q., Li, R., Wang, Q., & Chen, S. (2021). Non-fungible token (NFT): Overview, evaluation, opportunities and challenges. arXiv preprint arXiv:2105.07447.

ANEXO V: OPTIMIZATION OF SIEM – SECURITY INFORMATION EVENT MANAGEMENT USING DLP – DATA LOSS PREVENTION AND PANN – PARACONSISTENT ARTIFICIAL NEURAL NETWORK

Ano	Artigo	Status
2023	*OPTIMIZATION OF SIEM – SECURITY INFORMATION EVENT MANAGEMENT USING DLP – DATA LOSS PREVENTION AND PANN – PARACONSISTENT ARTIFICIAL NEURAL NETWORK*	Liliam Sayuri Sakamoto, Jair Minoro Abe, Luiz Antonio de Lima.
	Submetido para a Revista A1	*IEEE Systems Journal*
	SiteScore	9.1
	Highest Percentile	89% 84/792 Computer Science Applications



Submission Confirmation

Print

Thank you for your submission

Submitted to IEEE Systems Journal

Manuscript ID ISJ-RE-23-17138

Title OPTIMIZATION OF SIEM – SECURITY INFORMATION EVENT MANAGEMENT USING DLP – DATA LOSS PREVENTION AND PANN – PARACONSISTENT ARTIFICIAL NEURAL NETWORK

Authors SAKAMOTO, LILIAM
de Lima, Luiz
Abe, Jair

Date Submitted 12-Oct-2023

OPTIMIZATION OF SIEM – SECURITY INFORMATION EVENT MANAGEMENT USING DLP – DATA LOSS PREVENTION AND PANN – PARACONSISTENT ARTIFICIAL NEURAL NETWORK

Liliam Sayuri Sakamoto[1[0000-0001-8636-0100]], Jair Minoro Abe[2[0000-0003-2088-9065]], Luiz Antonio de Lima[3[0000-0003-4228-2387]]

[1,2,3,] Paulista University, 1212 Dr. Bacelar Street, São Paulo, Brazil.

liliam.sakamoto@aluno.unip.br

ABSTRACT

All companies suffer from cyberattacks by hackers, regardless of their size, so most of them try to protect their customers' data, which are primarily sensitive and subject to possible digital crimes with their malicious use. This study presents the improvement of a SIEM - Security Information Event Monitoring, which monitors Antivirus data, with the implementation of the DLP - Data Loss Prevention tool in conjunction with a Paraconsistent Artificial Neural Network (PANN) to bring greater assertiveness in the analysis of alerts that are dismissed as inconclusive. The methodology in its first stage is based on bibliographical research based on the three cycles of design for scientific research. In the second stage, it presents exploratory research with tests in an anonymized database of cybersecurity alerts (200 cases) of a transport company for 14 months extracted from a SIEM focused on Antivirus. At this stage, the study suggests implementing a

DLP - Data Loss Prevention developed in a computational program in the Python language, with an artificial intelligence technique focused on Deep Learning within one of the layers of the Paraconsistent Artificial Neural Network. The results show that in the pure analysis of the SIEM, 37% of the alerts considered as inconclusive data were discarded; however, with the application of the suggested solution, it is possible to reach the minimization of discard of only 20% of cybersecurity alerts, representing an optimization of 17%.

Keywords: Data Loss Prevention - DLP, Paraconsistent Artificial Neural Network - PANN, Security Information Event Monitoring – SIEM.

.

1.INTRODUCTION

This article aims at optimizing decision-making regarding internal and external alerts of a SIEM tool – Security Information and Event Monitoring, which monitors and centralizes alerts coming from other Host Antivirus applications, that is, from endpoints (equipment end users) and corporate servers (which can be on-premises or in the cloud). These alerts may indicate cybersecurity risks, such as port scan (malicious scanning in order to gain improper access), ransomware (data hijacking), phishing (social engineering in order to capture data through false emails), a scammer (scam on the Internet), malware (a virus that contaminates the environment), among others, which can have a significant impact on an organization and sanctions due to LGPD, as well as the improper disclosure of personal data of customers and employees, (NIST, 2023).

The first part of the study was based on bibliographical research adapted from Hevner (2007) that refers to three design cycles: the cycle of relevance, the cycle of rigour, and the central design cycle.

The SIEM tool captured the data used in the exploratory research part of two market applications (Trend Micro, 2023): Apex One and Deep Security, both developed by Trend Micro.

The database captured in this study presents a universe of 200 alerts, from which only the percentage of 37% were discarded as inconclusive alerts were selected, that is, 74 alerts. A DLP tested these - Data Loss Prevention developed in Python with a Paraconsistent Artificial Neural Network (PANN).

In the market, several processes use automated market tools. We will use Apex One Trend Micro and Microsoft Purview. However, these do a screening of raw data, but when detecting suspicious situations, the evaluation needs the opinion of a data engineer, that is, a specialist in the field of cybersecurity and also data protection and privacy, with a focus on LGPD.

The alerts are for decision-making in response to cybersecurity and data privacy incidents, which may prevent the spread of ransomware, greyware, malware, or phishing through spam before they contaminate the entire company or its data is kidnapped.

This study presents exploratory research with a qualitative approach (Souza et al., 2011) and the use of DLP – Data Loss Prevention and Paraconsistent Artificial Neural Network (PANN), for making assertive decisions about the detection of alerts and vulnerabilities (De Lima, 2021) referring to cybersecurity and privacy of personal data (LGPD, 2018).

1.1. SIEM

First, it should be clarified what a SIEM - Security Information and Event Management or Information Security System for Event Monitoring is. As the name implies, it refers to a set of tools and services that manage to monitor situations related to alerts. These alerts can be classified

as low, medium, or high risk and may be false positives. Companies worldwide are investing in information security (Weishäupl, 2018).

The term SIEM, or Security Information and Event Management, was first mentioned in a report by Gartner Inc. (Williams & Nicolett 2005).

Security Information and Event Management – SIEM is a simplification composed of two terms: Security Information Management (SIM) and Security Event Management (SEM) (Goldstein et al., 2013).

In 2011, the replacement of BI – Business Intelligence by SIEMs that had already stood out for a decade was already evident (Lozito, 2011) — addressing the importance of using SIEMs in companies for security monitoring, with the increasing use of business transactions worldwide. This use has intensified even more after the advent of the Covid-19 pandemic.

However, recently, with the occurrence of the war in Eastern Europe between Russia and Ukraine, spam attacks to spread ransomware, which kidnaps data in exchange for bitcoins, are flooding companies around the world, as the intention of these groups is to try to monetize and raise money for war maintenance — noting that the group of Russian cyber terrorists is already known to be constantly focused on monetization actions, according to research presented by the company Trend Micro, a specialist in antivirus tools (Trend Micro, 2023), which with its tool intercepted in 2022: 79,945,411,146 email threats.

There are ready-made market tools, which we call software packages (the best known Splunk), which can be used by companies depending on their size in conjunction with the SOC service - Security Operations Centers or Security Operations Center, which generally operate in a 24x7 situation, with teams on duty monitoring the environment at all times. However, not all companies have this need. Depending on the niche in which they operate, some end their activities, such as the cash transport logistics companies, at 6:00 p.m. and resume only at 6:00 a.m. due to the high business risk at night.

According to Lozito (2011), most SIEMs only collect logs of:

• Firewall traffic;

• IPS or IDS events;

• VP Events; It is

• Host.

SIEMs were deployed to prevent, detect and react against cyberattacks and provide visibility to identify high-risk areas and help with mitigation strategy, reduced costs, and incident response time. Most converge in BIG DATA systems (González-Granadillo et al., 2021).

SIEM is also considered an emerging technology in the cybersecurity area to normalize and correlate a large number of security events as if it were a single system in an ecosystem (Grammatikis, 2021).

SIEMs are essential tools used by Security Operations Centers (SOC) and Computer Security Incident Response Teams (CSIRT), with them correlated events that undergo forensic analysis and need to be identified as malicious or not in real-time. These teams face many challenges daily in the face of the increasing number of cyberattacks (Bhatt, 2014).

There are proposals for SIEMs with a sociotechnical structure to support organizations in the maturity of security risk escalation and support SOCs in overcoming their limitations (AlSabbagh & Kowalski, 2016).

SIEM can also use various non-standardized logs for security analysis (Cinque et al., 2018). This study is an exploratory analysis aimed at complementing current SIEM practice.

When market SIEMs are used, ready-made solutions, although pre-configured, users need to adjust the solution for each context (Kavanagh et al., 2015).

As SIEMs help to understand large amounts of data for security monitoring, noting that visualization is essential, it is necessary to apply human perception in the analysis of information (Novikova & Kotenko, 2013).

The main features of a SIEM are the collection and management of log events, real-time threat detection, compliance, connection to Active Director - AD, firewall monitoring, and presentation in dashboards and reports related to incidents (Gartner, 2018).

1.2. DLP – Data Loss Prevention

The concept of DLP – Data Loss Prevention is not new, but it is being used more nowadays, as ready-made tools on the market already have some functionality of this built-in monitoring standard. Applications such as Microsoft Office 365 can already bring some standardized models internally for compliance with the LGPD or other regulations such as GDPR. Data loss may occur when trying to carry out an analysis for compliance with the standards, with the lack of depth in the evaluation of the security criteria [15].

Trend Micro's Antivirus features a solution called Apex One, which is a SaaS – System as a Service for endpoint and server security, with the functionality to report logs on DLP – Data Loss Prevention, which can be configured according to the company's needs, for example, search for disclosure of information on CPF, CNPJ, Invoices, credit card, social security, NFT or even specific words, and others. (Trend Micro, 2023)

Each of these tools uses artificial intelligence algorithms that cross data and monitor patterns pre-formatted by the information security analyst, starting from classical logic for their conclusions and presentations for decision-making, mainly when data loss, theft or leakage is detected.

The DLP must have an assertive configuration to capture results in line with the company's needs, such as the number of credit card transactions per customer, which ones are valid, or if there are blockages for the same card, among other situations, to assess suspicious situations [16].

DLP can also be standardized to evaluate suspected malware, ransomware, and spam using public safety lists such as the Miter Att&ck website (Miter, 2023).

Many companies have developed tools that manage to unify alerts about malicious activities such as IP scanning, data flow, malicious emails, intrusion attempts, firewall monitoring, review of security policy updates, and implemented security patches (Nist, 2023).

DLP tools require testing and calibration, in which alerts with excessive false positives can occur, but they use classic logic (De Lima, 2021).

The sophistication reached by malicious software requires constant efforts to mitigate this practice with preventive solutions, such as using DLP for effective monitoring in conjunction with Paraconsistent Logic, to innovate and be one step ahead of them.[18].

1.3. Artificial Intelligence and Artificial neural networks

The artificial intelligence technique was proposed in the 1940s by Warren McCulloch and Walter Pitts (1943), currently 80 years old, since these researches were based on the philosophy, knowledge, and function of neurons. Whereas Russell and Whitehead created propositional logic. Another pillar is the theory of Turing computation.

These two researchers McCulloch and Pitts proposed a model of artificial neurons, where each neuron can be characterized by being “on” or “off”, and when switching to “on” there is stimulation by a sufficient number of neighboring neurons.

The state of a neuron can be considered “equivalent in concrete terms to a proposition that defined its stimulus”. Examples show that any computable function can be calculated by a certain network of neurons connected in all kinds of structural logic connectors (and, or, not, etc.). McCulloch and Pitts also suggested that when networks were assertively defined, they would be able to learn (De Lima et al, 2021).

The paraconsistent neuron is proposed to serve the paraconsistent neural network (De Lima et al, 2021). Another relevant aspect within each context is that the neuron chooses which characteristic (x) will be used in the network to be trained.

Figure 1. Paraconsistent Artificial Neutron. Source: De Lima, 2021.

The use of Artificial Neural Networks is mainly intended for the treatment of significant concepts such as uncertainty, inconsistency (contradiction), and paracompleteness, going against the Annotated Paraconsistent Logic Evidential Et (Abe et al, 2008).

The architecture of Deep Learning or deep Learning can be verified in the form of recurrent neural networks and deep convolutional neural networks so that they can be applied to areas of alert identification, recognition of framework patterns such as NIST and Miter Attack, processing of SIEMs logs and Cybersecurity (De Lima et al., 2021).

It allows the configuration of computational models composed of several layers of processing, and these can learn data representations at various levels of abstraction. This method also uses the backpropagation algorithm to indicate how changes should be made to its internal parameters to calculate each layer’s representation from the previous layer’s representation (LeCun et al., 2015). It is understood that the qualification of holding something is a human behaviour that is also shown as an essential factor of machines today (Shinde & Shah, 2018).

Deep Learning is a technique used based on artificial neural networks. In recent years it has become a powerful machine learning tool to reshape artificial intelligence, with rapid improvements in computational power, faster data storage, and parallelization. Another factor that led to this technology’s immediate acceptance is its predictive power, the ability to generate high-level optimized features automatically, and detailed interpretation of input data (Ravi et al, 2016).

Deep Learning, despite having initial concepts in the 80s, is focused on the characterization of neural networks, where it presents the Artificial Neural Network notation that does not consider the input layer, that is, according to Figure xx, we have three layers (Layer 1, Hidden layer and Layer 2). In this sense, Paraconsistent Artificial Neural Networks (PANN) are proposed in a similar way.

In addition to considering it as a hidden layer for not knowing what is being used in the training set. In Figure 3b we have three layers (Layer 1, Hidden Layer, and Layer 2).

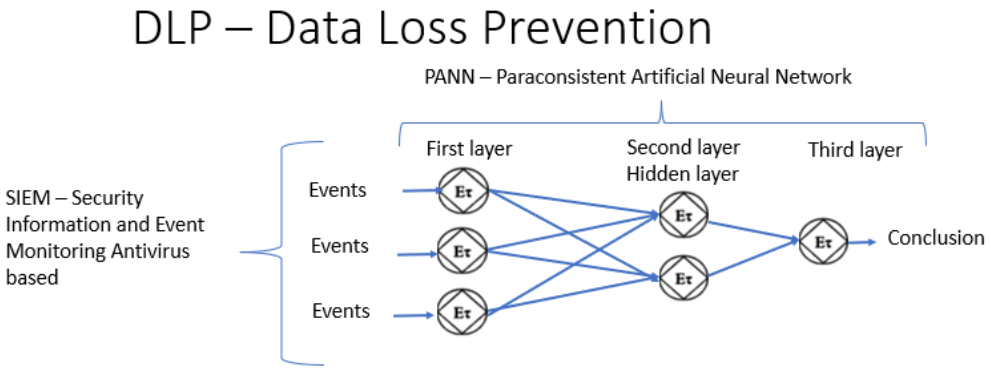


Figure 2. DLP – Data Loss Prevention. Source: Authors.

1.4. Paraconsistent Artificial Neural Network - PANN

In 2004, Abe presented the theory called Paraconsistent Artificial Neural Network – PANN, built based on the para-analyzer algorithm, with the ability to analyze and manipulate data considered uncertain, inconsistent, and paracomplete (Abe & Nakamatsu, 2012).

Paraconsistent Artificial Neural Network (PANN) is an appropriate tool to solve problems such as pattern prediction and recognition, with the ability to deal directly with imprecise data, providing quantitative and qualitative analysis of these (da Silva Lopes et al., 2010).

For the medical area, when using the Paraconsistent Artificial Neural Network, it is possible to determine the degree of certainty of a diagnosis, going deeper until reaching a Kappa index at a rate of 80%. (Abe, Prado & Nakamatsu, 2006; Abe et al, 2007; Abe et al, 2010; Abe & Lopes, 2011; Abe et al, 2012; Abe, Lopes & Anghinah, 2015). It helps to mitigate the effects of false positive and true positive diagnoses (Do Amaral et al., 2015). Other applications have also been carried out in the areas of orthodontics (Mario et al., 2010; Bahaa, 2011) and biomedicine (Abe et al., 2013).

PANN can be used to trigger system alarms based on dynamic monitoring (da Cruz et al., 2015).

The basis of Paraconsistent Artificial Neural Networks lies in their intrinsic ability to deal with imprecise, inconsistent, and paracomplete data (Abe et al., 2005; Abe, Lopes & Nakamatsu, 2008; Da Silva Filho, 2010; Souza et al., 2013; Abe et al. al, 2018).

.The recognition of computational patterns by a PANN demonstrates the importance of Artificial Intelligence tools present in numerous areas of knowledge with applications in various topics (Souza & Abe, 2015). The PANN methodology can work with inaccurate, inconsistent, and paracomplete data without being trivial (Minoro Abe et al., 2012; Souza, 2014).

Non-classical logic has become a powerful tool to aid decision-making. This type of logic highlights the clarity of containing provisions contrary to some of the basic principles of Aristotelian logic, such as the principle of contradiction. (De Lima et al, 2020).

Within the structure of the PANN, Paraconsistent Artificial Neural Cells are built that can learn specific signals in the form of functions applied to their inputs (Ferrara et al., 2005).

A PANN can be used to build an efficient architecture to determine and monitor quality indices with applications in several areas of engineering (da Silva Filho et al., 2016).

Solutions such as PANN using Artificial Intelligence in agriculture can contribute to Brazilian agribusiness, as it is still unexplored (de Souza et al., 2021).

Through tests with the PANN architecture, expert systems can configure a profile for quantifying grouped information groups (Abe, Lopes & Nakamatsu, 2012).

Comparisons between a model elaborated in PANN and the performance of the human expert provided kappa indices that varied from moderate to almost perfect agreement — being the concordance between the performance of the PANN and the equivalent expert. Contradictions presented in the data that the experts did not notice were pointed out, which highlights the contribution of this type of system in supporting decision-making (Mario et al., 2010).

Studies have shown that a PANnet demonstrates dynamic properties with robustness to perturbations, both in the learning and imitation processes (Ribeiro, 2008; Da Silva Filho, 2023).

Artificial neural networks have been used in recent decades to perform tasks with comparative Learning, where input-output models must be adapted, identify a given system, and control and recognize patterns involving uncertain situations that can be challenging (de Carvalho et al., 2021).

Assuming that it is a type of deep neural network that mainly uses the Paraconsistent annotated evidential logic Eτ, the Paraconsistent Artificial Neural Network (PANN) is internally constituted by the Paraconsistent Artificial Neural Cells family.

In this study, only the Paraconsistent Artificial Neural Cell of Passage and Decision (CNAPpd) will be detailed, which works as a signalling of levels with maximum or minimum signals. When comparing the input signal and the Decision Factor (FtD) value, the signal is sent with the representation of True or undefined to the output. Therefore, this cell will only display two values (Abe et al., 2008):

- value 1, representing the status “true” and
- value 0.5, representing the “undefined” status.

Represented by the following algorithm for the CNAPpd:

Input variables: μ1, such that: $0 \leq \mu 1 \leq 1$.

The complement of the input variable is $\lambda = 1 - \mu 1$

The degree of Evidence is calculated by:

$$\mu_E = \frac{(\mu_1 - \lambda) + 1}{2}$$

The limit values of False and True:

$$Vl_V = \frac{1 + FT_D}{2} \; e \; Vl_F = \frac{1 - Ft_D}{2}$$

If $Vl_F \geq \mu_{ER} \geq Vl_V$ so: $S_1 = 1$ */True*/

If not, $S_1 = 0{,}5$ */Indefinition*/

CNAPpd uses the input and output variables with the adjustment signals.

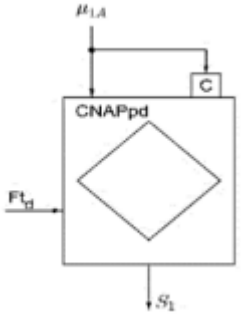


Figure 3: Passing and Decisioning Paraconsistent Artificial Neural Cell (CNAPpd). Source: Abe et al., 2008.

The Passage and Decision Paraconsistent Artificial Neural Cell (CNAPpd) is a Paraconsistent Artificial Neural Unit – UNAP, which stands out for allowing the treatment of extremes (Table 3) and non-extreme states (Table 4).

Table 1 – Extreme States Source: Abe et al. (2011).

Extreme States	**Symbol**
True	V
False	F
Inconsistent	T
Paracomplete	⊥

Table 2 provides the expert with analysis and support that can be adjusted to plausible levels during the analysis.

Table 2. Non-extreme states

Non-extreme states	**Symbol**
Quasi**-true tending to Inconsistent**	QV⟶T
Quasi-true tending to Paracomplete	QV⟶⊥
Quasi-false tending to Inconsistent	QF⟶T
Quasi-false tending to Paracomplete	QF⟶⊥
Quasi-inconsistent tending to True	QT⟶V
Quasi-inconsistent tending to False	QT⟶ F
Quasi-paracomplete tending to True	Q⊥⟶V
Quasi-paracomplete tending to False	Q⊥⟶ F

2. Methodology

This article was based on the Design Science Research paradigm (Hevner, 2007), which is:

- o Relevance Cycle: The contextual environment of this research is the cyber security of corporate environments based on the LGPD, and for the exploratory research (anonymized), alerts were extracted from a SIEM focused on the Antivirus of a Carrier company for 14 months for testing.
- Cycle of Rigor: We use a DLP – Data Loss Prevention elaborated in a Python Program together with a Paraconsistent Artificial Neural Network – PANN to minimize the amount of data discarded from the SIEM analysis.

- The core design cycle supports the continuous improvement of the research process focused on Artificial Intelligence techniques focused on Deep Learning located within the PANN at Level 2 - the hidden layer.

It also composes exploratory research with an anonymized database used by a transport company that requested confidentiality of 14 months extracted from a SIEM - Security Information and Event Monitoring focused only on Antivirus files.

A still picture of these data is presented and compared afterwards with the tested data.

These data undergo the analysis of the DLP - Data Loss Prevention developed in a computational algorithm composed of a Paraconsistent Artificial Neural Network (PANN), which has the following layers:

Layer 1:

Captures events extracted from SIEM focused on Antivirus (all types of alerts).

Layer 2 (hidden layer):

Application of the Artificial Intelligence technique with a focus on Deep Learning for training the paraconsistent artificial neuron (figure 2).

When using the technique of artificial intelligence (training, artificial neuron) (Russel, 2010), the neuron can be improved by learning from the data, which specific improvements can be used to make them, but which depend on four main factors :

- The component that should be improved is the paraconsistent artificial neuron.
- The paraconsistent artificial neuron presents prior knowledge of some automated detection patterns.
- The representation is based on the Miter Attack framework standard, a component of the Antivirus tool for hosts (endpoints and servers) from which the alert logs were previously extracted.
- The feedback that the paraconsistent artificial neuron must learn is to minimize the information discarded as inconclusive.

Layer 3:

In this layer, the result of the para-analyzer is reached, applying the Annotated Evidential Et Et Paraconsistent Logic to treat data that were considered inconclusive in the previous layer, resulting in the four extreme states (table 2).

3. Results and discussions

Anonymized data collected from the Carrier company, extracted from the SIEM, 74 suspicious alerts were discarded, that is, 37% of a total of 200 alerts:

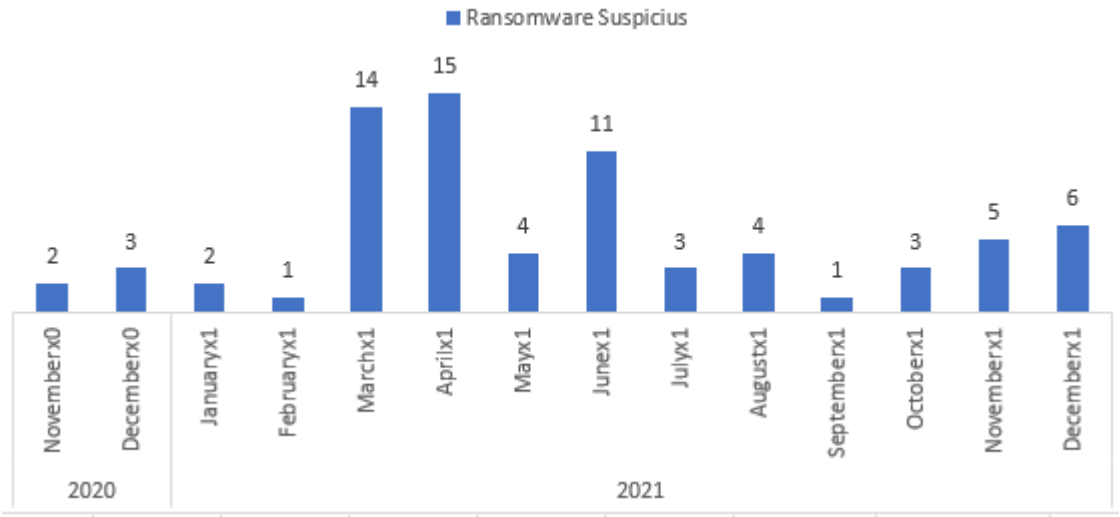


Figure 4. SIEM Report – Suspicious ransomware alerts were discarded. Source: Authors.

The Python program passes data:

```
import csv
from typing import List
from util import normalize_by_feature_scaling
from network import Network
from random import shuffle

if __name__ == "__main__":
    dlp_parameters: List[List[float]] = []
    dlp_classifications: List[List[float]] = []
    dlp_species: List[int] = []
    with open(dlp.csv, mode='r') as dlp_file:
        dlp: List = list(csv.reader(dlp of data in random order))
        for dlp in nfts:
            parameters: List[float] = [float(n) for n in dlp[1:14]]
            dlp_parameters.append(parameters)
            species: int = int(dlp[0])
            if species == 1:
                dlp_classifications.append([1.0, 0.0, 0.0])
            elif species == 2:
                dlp_classifications.append([0.0, 1.0, 0.0])
            else:
                dlp_classifications.append([0.0, 0.0, 1.0])
            dlp_species.append(species)
    normalize_by_feature_scaling(nft_parameters)
# Lista [camada de entrada,camada intermediaria, camada de saida], taxa de Aprendizag
    dlp_network: Network = Network([13, 7, 3], 0.9)

    def dlp_interpret_output(output: List[float]) -> int:
        if max(output) == output[0]:
            return 1
        elif max(output) == output[1]:
            return 2
        else:
            return 3
    # train over the first 150 dlp 10 times
    dlp_trainers: List[List[float]] = dlp_parameters[0:150]
    dlp_trainers_corrects: List[List[float]] = dlp_classifications[0:150]
    for _ in range(10):
        dlp_network.train(dlp_trainers, dlp_trainers_corrects)

    # test over the last 28 of the dlp in the data set
    dlp_testers: List[List[float]] = dlp_parameters[150:178]
    dlp_testers_corrects: List[int] = dlp_species[150:178]
    dlp_results = dlp_network.validate(dlp_testers, dlp_testers_corrects, dlp_interpr
    print(f"{dlp_results[0]} correct of {dlp_results[1]} = {dlp_results[2] * 100}%")
```

Figure 5. Python Program. Source: Authors.

The result after tests with the PANN on top of the data discarded as suspicious, 74 data were analyzed, where 54% of data were considered abnormal and validated as real risk situations (“true”) and 46% validated as false positives (“undefined”):

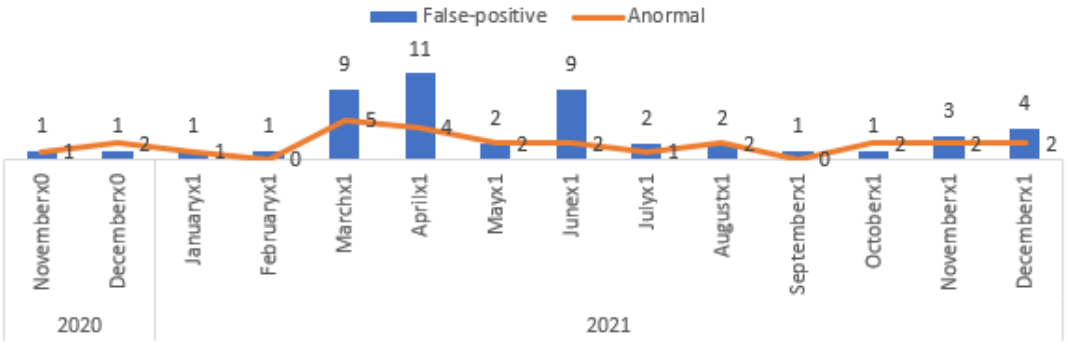


Figure 6. DLP Report. Source: Authors.

Studies show that several companies use the SIEM as a log and incident information concentrator. However, the number of alerts and information that are discarded is countless due to the lack of sufficient specialized labour for detailed analysis and because they are considered inconsistent.

In this article, we highlight the results of the approach of artificial intelligence techniques using DLP and the Paraconsistent Artificial Neural Network in optimizing decision-making regarding the definition of incident response actions and avoiding possible unidentified risks.

The neural architecture found in the family of paraconsistent neural networks used in this study was the Passage and Decision Paraconsistent Artificial Neural Cell (CNAPpd), which allowed gains by proposing possible solutions for responses to incident alerts.

4. Conclusion

It is concluded that the alignment between the use of DLP - Data Loss Prevention with a Paraconsistent Artificial Neural Network presents a significant gain for the issue of monitoring and analysis of the loss of cybersecurity alerts of a SIEM.

The implementation of this type of tool increased the performance of the analysis of the alerts concerning the SIEM, which disregarded 37% of the data as inconclusive to only 20% after the DLP together with the Paraconsistent Neural Network - PANN, considering that even these intervals can be monitored so that its 100% effectiveness of the environment data.

It is understood that the union of the knowledge of data privacy professionals with the improvement of tools such as DLP together with the Paraconsistent Neural Network - PANN, brings the possibility of significant advances in the adequacy of the LGPD and increased assertiveness in the identification of decision making beyond of SIEM, for cybersecurity within the corporate environment.

ACKNOWLEDGMENTS

This study was partially financed by the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financial Code 001.

References:

Abe, J. M., Lopes, H. F. D. S., & Anghinah, R. (2007). Paraconsistent artificial neural networks and Alzheimer disease: A preliminary study. Dementia & Neuropsychologia, 1, 241-247.

Abe, J. M., Lopes, H. F., & Nakamatsu, K. (2013). Paraconsistent artificial neural networks and EEG. International Journal of Knowledge-based and Intelligent Engineering Systems, 17(2), 99-111.

Abe, J. M., Lopes, H. F., Nakamatsu, K., & Akama, S. (2010). Paraconsistent artificial neural networks and EEG analysis. In Knowledge-Based and Intelligent Information and Engineering Systems: 14th International Conference, KES 2010, Cardiff, UK, September 8-10, 2010, Proceedings, Part III 14 (pp. 164-173). Springer Berlin Heidelberg.

Abe, J. M., Lopes, H. F., Nakamatsu, K., & Akama, S. (2011). Applications of paraconsistent artificial neural networks in EEG. In Computational Collective Intelligence. Technologies and Applications: Third International Conference, ICCCI 2011, Gdynia, Poland, September 21-23, 2011, Proceedings, Part I 3 (pp. 82-92). Springer Berlin Heidelberg.

Abe, J. M., Lopes, H. F., & Nakamatsu, K. (2012). Paraconsistent artificial neural networks and delta, theta, alpha, and beta bands detection. Advances in Reasoning-Based Image Processing Intelligent Systems: Conventional and Intelligent Paradigms, 331-364.

Abe, J. M., Prado, J. C. A., & Nakamatsu, K. (2006). Paraconsistent artificial neural network: applicability in computer analysis of speech productions. In Knowledge-Based Intelligent Information and Engineering Systems: 10th International Conference, KES 2006, Bournemouth, UK, October 9-11, 2006. Proceedings, Part II 10 (pp. 844-850). Springer Berlin Heidelberg.

Abe, J. M., Ortega, N. R., Mário, M. C., & Del Santo Jr, M. (2005, September). Paraconsistent artificial neural network: An application in cephalometric analysis. In International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (pp. 716-723). Berlin, Heidelberg: Springer Berlin Heidelberg.

Abe, J. M., Lopes, H. F., & Nakamatsu, K. (2008). Improving EEG analysis by using paraconsistent artificial neural networks. In Knowledge-Based Intelligent Information and Engineering

Abe, J. M., & Nakamatsu, K. (2012). Paraconsistent artificial neural networks and pattern recognition: Speech production recognition and cephalometric analysis. Advances in Reasoning-Based Image Processing Intelligent Systems: Conventional and Intelligent Paradigms, 365-382.

Abe, J. M., Lopes, H. F. S., & Nakamatsu, K. (2012). Paraconsistent artificial neural networks and AD analysis–Improvements. In Computational Collective Intelligence. Technologies and Applications: 4th International Conference, ICCCI 2012, Ho Chi Minh City, Vietnam, November 28-30, 2012, Proceedings, Part I 4 (pp. 259-267). Springer Berlin Heidelberg.

Abe, J. M., & Nakamatsu, K. (2013). PARACONSISTENT ARTIFICIAL NEURAL NETWORKS AND APPLICATIONS. In The Handbook on Reasoning-Based Intelligent Systems (pp. 307-330).

Abe, J. M., da Silva Lopes, H. F., & Anghinah, R. (2015). Paraconsistent neurocomputing and biological signals analysis. Paraconsistent Intelligent-Based Systems: New Trends in the Applications of Paraconsistency, 273-306.

Abe, J. M., Nakamatsu, K., Akama, S., & Filho, J. I. (2018). The Importance of Paraconsistency and Paracompleteness in Intelligent Systems. In Intelligent Decision Technologies 2017: Proceedings of the 9th KES International Conference on Intelligent Decision Technologies (KES-IDT 2017)–Part II 9 (pp. 196-205). Springer International Publishing.

Abe, Jair Minoro; da Silva Filho, João Inácio; Torres, Germano Lambert. Inteligência Artificial com as Redes de Análises Paraconsistentes. Teoria e Aplicações, Editora : LTC; 1ª edição, 314pag., 2008. ISBN-10 : 8521616317, ISBN-13 : 978-8521616313

Abe, J. M. (2004). Paraconsistent artificial neural networks: An introduction. In Knowledge-Based Intelligent Information and Engineering Systems: 8th International Conference, KES 2004, Wellington, New Zealand, September 20-25, 2004, Proceedings, Part II 8 (pp. 942-948). Springer Berlin Heidelberg.

Amrit T. Williams and Mark Nicolett. 2005. Improve IT Security With Vulnerability Management. Technical Report - Gartner Inc. (2005)

Bahaa, K., Noor, G., & Yousif, Y. (2011). The artificial intelligence approach for diagnosis, treatment and modelling in orthodontic. Principles in contemporary orthodontics.

Bhatt, S., Manadhata, P. K., & Zomlot, L. (2014). The operational role of security information and event management systems. IEEE security & privacy, 12(5), 35-41.

Cinque, M., Cotroneo, D., & Pecchia, A. (2018, October). Challenges and directions in security information and event management (SIEM). In 2018 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW) (pp. 95-99). IEEE.

da Cruz, C. M., Rocco, A., Mario, M. C., Garcia, D., Lambert-Torres, G., Abe, J. M., ... & da Silva Filho, J. I. (2015, September). Application of paraconsistent artificial neural network in statistical process control acting on voltage level monitoring in electrical power systems. In 2015 18th international conference on intelligent system application to power systems (ISAP) (pp. 1-6). IEEE.

da Cruz, C. M., Rocco, A., Mario, M. C., Garcia, D., Lambert-Torres, G., Abe, J. M., ... & da Silva Filho, J. I. (2015, September). Application of paraconsistent artificial neural network in statistical process control acting on voltage level monitoring in electrical power systems. In 2015 18th international conference on intelligent system application to power systems (ISAP) (pp. 1-6). IEEE.

Da Silva Filho, J. I., Fernandes, C. L. M., Silveira, R. S. D., Gomes, P. M., Matos, S. L. D. C., Santo, L. D. E., ... & Lambert-Torres, G. (2023). Process of Learning from Demonstration with Paraconsistent Artificial Neural Cells for Application in Linear Cartesian Robots. Robotics, 12(3), 69.

Da Silva Filho, J. I., Torres, G. L., & Abe, J. M. (2010). Uncertainty treatment using paraconsistent logic: Introducing paraconsistent artificial neural networks (Vol. 211). IOS Press.

da Silva Filho, J. I., Da Cruz, C. M., Rocco, A., Garcia, D. V., Ferrara, L. F. P., Onuki, A. S., ... & Abe, J. M. (2016). Paraconsistent artificial neural network for structuring statistical process control in electrical engineering. In Towards Paraconsistent Engineering (pp. 77-102). Cham: Springer International Publishing.

da Silva Lopes, H. F., Abe, J. M., & Anghinah, R. (2010). Application of paraconsistent artificial neural networks as a method of aid in the diagnosis of Alzheimer disease. Journal of medical systems, 34, 1073-1081.

de Lima, L. A., Abe, J. M., Martinez, A. A. G., de Frederico, A. C., Nakamatsu, K., & Santos, J. (2020). Process and subprocess studies to implement the paraconsistent artificial neural networks for decision-making. In Recent Trends in Intelligent Computing, Communication and Devices: Proceedings of ICCD 2018 (pp. 503-512). Springer Singapore.

de Lima, L. A., Abe, J. M., Martinez, A. A. G., de Souza, J. S., Bernardini, F. A., de Souza, N. A., & Sakamoto, L. S. (2021). Study of Pann Components in Image Treatment for Medical Diagnostic Decision-Making/Estudo De Componentes Pann no Tratamento de Imagem para Tomada de Decisão de Diagnóstico Médico. Revista FSA (Centro Universitário Santo Agostinho), 18(7), 173-186.

de Lima, L. A., Abe, J. M., Martinez, A. A., Sakamoto, L. S., & de Lima, L. P. (2021). Application of architecture using AI in the training of a set of pixels of the image at aid decision-making diagnostic cancer. Procedia Computer Science, 192, 1740-1749.

do Amaral, F. V., Abe, J. M., Cadim, A. J. S., Kirilo, C. Z., Baltazar, C. A., Pereira, F. L., ... & Oliveira, C. C. (2015). Paraconsistent artificial neural network applied in breast cancer diagnosis support. In Advances in Production Management Systems: Innovative Production Management Towards Sustainable Growth: IFIP WG 5.7 International Conference, APMS 2015, Tokyo, Japan, September 7-9, 2015, Proceedings, Part I 0 (pp. 464-472). Springer International Publishing.

de Souza, T. T. A., de Oliveira, C. C., Abe, J. M., Aharari, A., & Nakamatsu, K. (2021). Paraconsistent artificial neural network applied to agribusiness. In New Developments of IT, IoT and ICT Applied to Agriculture: Proceedings of ICAIT 2019 (pp. 19-28). Springer Singapore.

Ferrara, L. F. P., Yamanaka, K., & da Silva Filho, J. I. (2005, May). A system of recognition of characters based on paraconsistent artificial neural networks. In Proceedings of the 2005 conference on Advances in Logic Based Intelligent Systems: Selected Papers of LAPTEC 2005 (pp. 127-134).

Gartner Inc. 2018. Security Information and Event Management (SIEM). (2018). Retrieved 20.06.2018 from https://www.gartner.com/en/information-technology/glossary/security-information-and-event-management-siem

González-Granadillo, G., González-Zarzosa, S., & Diaz, R. (2021). Security information and event management (SIEM): analysis, trends, and usage in critical infrastructures. Sensors, 21(14), 4759.

Hevner, A. R. (2007). A three cycle view of design science research. Scandinavian journal of information systems, 19(2), 4.

Kavanagh, K. M., Rochford, O., & Bussa, T. (2015). Magic quadrant for security information and event management. Gartner Group Research Note.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. nature, 521(7553), 436-444.

Lozito, Ken. Mitigating risk: Analysis of security information and event management. **International Journal of Business Intelligence Research (IJBIR)**, v. 2, n. 2, p. 67-75, 2011.

Mario, M. C., Abe, J. M., Ortega, N. R., & Del Santo Jr, M. (2010). Paraconsistent artificial neural network as auxiliary in cephalometric diagnosis. Artificial Organs, 34(7), E215-E221.

Markus Goldstein, Stefan Asanger, Matthias Reif, and Andrew Hutchison. 2013. Enhancing Security Event Management Systems with Unsupervised Anomaly Detection. In ICPRAM. 530–538.

Mckinney, W. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, 1 edition. Beijing: O'Reilly Media, 2012

Minoro Abe, J., Lopes, H. F., & Nakamatsu, K. (2012). An Overview of Paraconsistent Artificial Neural Networks and Applications. Advances in Knowledge-Based and Intelligent Information and Engineering Systems, 1350-1359.

Mitre , 2023. https://attack.mitre.org/

NIST, 2023. https://www.nist.gov/cyberframework

Radoglou-Grammatikis, P., Sarigiannidis, P., Iturbe, E., Rios, E., Martinez, S., Sarigiannidis, A., ... & Ramos, F. (2021). Spear siem: A security information and event management system for the smart grid. Computer Networks, 193, 108008.

Ravì, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2016). Deep Learning for health informatics. IEEE journal of biomedical and health informatics, 21(1), 4-21.

Ribeiro, C. H. C., & da Silva, A. A. (2008). Map Matching Based on Paraconsistent Artificial Neural Networks. IFAC Proceedings Volumes, 41(2), 14675-14680.

Russell, S. J. (2010). Artificial intelligence a modern approach. Pearson Education, Inc..

Shinde, P. P., & Shah, S. (2018, August). A review of machine learning and deep learning applications. In 2018 Fourth international conference on computing communication control and automation (ICCUBEA) (pp. 1-6). IEEE.

Souza, S., Abe, J. M., & Nakamatsu, K. (2013). MICR automated recognition based on paraconsistent artificial neural networks. Procedia Computer Science, 22, 1083-1091.

Souza, S., & Abe, J. M. (2015). Paraconsistent artificial neural networks and aspects of pattern recognition. Paraconsistent Intelligent-Based Systems: New Trends in the Applications of Paraconsistency, 207-231.

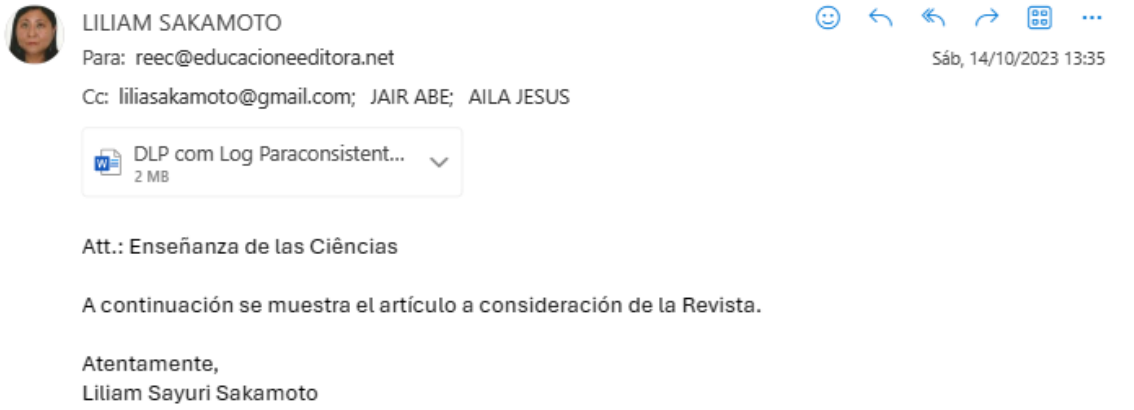
Souza, S., & Abe, J. M. (2014). Handwritten numerical character recognition based on paraconsistent artificial neural networks. In Recent Developments in Computational Collective Intelligence (pp. 93-102). Springer International Publishing.

Trend Micro. Disponível em: https://www.trendmicro.com/vinfo/br/security/research-and-analysis/threat-reports/roundup/annual-trend-micro-email-threats-report

ANEXO VI: DLP: prevención de pérdida de datos con lógica paraconsistente para la seguridad en el metaverso

Ano	Artigo	Status
2023	*DLP: prevención de pérdida de datos con lógica paraconsistente para la seguridad en el metaverso*	Liliam Sayuri Sakamoto, Jair Minoro Abe, Jonatas Santos de Souza, Luigi Pavarini de Lima, e Lucimara da Costa Santos Bernardini
	Submetido para Revista A1	*Enseñanzade las Ciencias*
	SiteScore	2.0
	Highest Percentile	52% 704/1469 Education

ISSN	Título	Área de Avaliação	Classificação
0212-4521	ENSEÑANZA DE LAS CIENCIAS	ENGENHARIAS III	A1



LILIAM SAKAMOTO

Para: reec@educacioneeditora.net

Sáb, 14/10/2023 13:35

Cc: liliasakamoto@gmail.com; JAIR ABE; AILA JESUS

DLP com Log Paraconsistent...
2 MB

Att.: Enseñanza de las Ciências

A continuación se muestra el artículo a consideración de la Revista.

Atentamente,
Liliam Sayuri Sakamoto

DLP: prevención de pérdida de datos con lógica paraconsistente para la seguridad en el metaverso

[a]Liliam Sayuri Sakamoto

[b]Jair Minoro Abe

[c]Jonatas Santos da Silva

[d]Luigi Pavarini de Lima

[e]Lucimara da Costa Santos Bernardini

[a] Paulista University, Brazil, [a] liliam.sakamoto@aluno.unip.br.

[b] Paulista University, Brazil, b jair.abe@docente.unip.br .

[c] Paulista University, Brazil, c jonatas1516@gmail.com .

[d] Institute of Mathematics and Statistics (IME) - University of São Paulo-USP, Brazil, [d] aula.prof.luiz@gmail.com

[e] Damásio de Jesus College, [e] lucimara_costasantos@hotmail.com

Resumen: El objetivo de este artículo es proponer una estructura de investigación para optimizar esta seguridad de los activos con NFT - Non-fungible Token con el uso de DLP - Data Loss Prevention y Paraconsistent Logic para la identificación no solo de forma preventiva, sino activa de la pérdida, robo, uso indebido y fuga de este tipo de activos durante su uso en el Metaverso. Se realizó una revisión bibliográfica sobre Metaverso, DLP, Lógica Paraconsistente, técnicas de Inteligencia Artificial, NFT y Protección de Datos con foco en la LGPD (Ley Brasileña de Protección de Datos), en conjunto con una investigación exploratoria. Con un DLP y una base de datos facilitada por la empresa de transporte con 200 artículos analizados. Se verificó que una cantidad importante de datos serían descartados en la primera etapa del proceso (37%), ya que no presentan una definición activa sobre el estado de estos bienes. considerando la creciente innovación tecnológica con el uso del Metaverso, como ambiente de interacción educativa, empresarial y gubernamental frente al riesgo de ataques cibernéticos, urge la necesidad de fortalecer su seguridad, más aún cuando en este ambiente, donde existe la posibilidad de mover activos con NFT que son objetos de gran valor adquiridos y transados en este medio. Con el uso del programa Python en el DLP se observó que presentó un 37% de pérdida de datos en su análisis con este proceso de Inteligencia Artificial solo con el desempeño del DLP, en comparación con la optimización de este análisis con el uso de Lógica Paraconsistente al 23%, es decir, un aprovechamiento de más del 15% de los datos.

Palabras clave: DLP – Data Loss Prevention, Lógica Paraconsistente Anotada Evidencial Eт, Inteligencia artificial, Metaverso, NFT – Tokens no fungibles.

TITLE: *DLP – DATA LOSS PREVENTION WITH PARACONSISTENT LOGIC ANNOTATED FOR SECURITY IN THE METAVERSE*

Abstract: The objective of this article is to propose a research structure to optimize the security of assets with NFT - Non-fungible Token with the use of DLP - Data Loss Prevention and Paraconsistent Logic for identification not only in a preventive way but actively the loss, theft, misuse, and leakage of these types of assets during use in the Metaverse. A bibliographic review was carried out on Metaverse, DLP, Paraconsistent Logic, Artificial Intelligence techniques, NFT, and Data Protection with a focus on LGPD (Brazilian Data Protection Law), in conjunction with exploratory research. With a DLP and a database provided by the transport company with 200 analyzed items. It was verified that a significant amount of data would be discarded in the first stage of the process (37%) since they do not present an operational definition regarding the status of these assets. considering the growing technological

innovation with the use of the Metaverse, as an environment for educational, corporate, and governmental interaction as opposed to the risk of cyber attacks, there is an urgent need to strengthen its security, especially when in this environment, where there is the possibility of moving assets with NFT which are objects of great value acquired and transacted in this medium. With the use of the Python program in the DLP, it was observed that it presented 37% of data loss in its analysis with this Artificial Intelligence process only with the performance of the DLP, in comparison with the optimization of this analysis with the use of Paraconsistent Logic to 23%, that is, a use of more than 15% of the data.

Keywords: DLP – Data Loss Prevention, Logic Paraconsistent Annotated Evidential Eᴛ, Artificial Intelligence, Metaverse, NFT - Non-fungible Tokens.

1. Introducción

1.1 Contexto General

Es interesante como la innovación tecnológica con el uso del entorno Metaverso, implementó la evolución en la captura de datos y en la transformación de los mismos en información válida como el uso de los NFT - Non-fungible Tokens. Hoy en día, los datos estructurados digitalmente se pueden identificar fácilmente en este token, ya que son únicos e individuales y se pueden almacenar dentro del servidor de base de datos de una estructura.

Sin embargo, cómo garantizar la seguridad de todos los activos con NFT válidos, sin evitar su pérdida, robo, fuga de estos activos, e incluso identificando aquellos con NFT no válidos, así como el tránsito de un punto a otro, dentro y fuera de la Metaverso de manera lícita.

El diferencial de cada organización en el seguimiento e identificación de estos activos radica en la capacidad de gestionar el riesgo de forma automatizada con tecnologías de inteligencia artificial, o como en esta propuesta con una implementación DLP, porque cuanto más se accede al Metaverso y se transforma su información con o sin la interacción del usuario, un activo puede presentar un valor creciente, por lo tanto, una mayor importancia en el aspecto de seguridad.

La aplicación de DLP - Data Loss Prevention delimita el flujo de estos activos con NFT monitoreando su entrada, si hay transformaciones, hasta su salida, donde el tránsito en el ambiente *Metaverse* debe ser seguro con la trazabilidad del *audit trail* y logs del programa en Python, ya que existe una monitorización activa, y que las alertas de pérdida, robo o fuga se pueden analizar mediante algoritmos de Inteligencia Artificial.

Si bien este tipo de herramienta es avanzada en contraste con el monitoreo manual diario, algunas alertas pueden presentar dudas, causando problemas en la trazabilidad. Por ello, la implementación de una capa más de análisis

con el Lógico Paraconsistente Anotado Evidencial Et añade mayor asertividad en la toma de decisiones.

Se realizó una investigación exploratoria con la recolección de una muestra de una empresa de transporte que solicita confidencialidad, ya que los datos fueron anonimizados de acuerdo con la LGPD - Ley General de Protección de Datos de Brasil (Brasil, 2022) para probar esta implementación, refiriéndose solo a un período de datos.

1.2 LGPD – Ley General de Protección de Datos

La protección de datos personales (Lima, L.A. de; et al, 2020), ya sea que se encuentren en empresas educativas, gubernamentales o privadas, que pueden ser o no datos sensibles, esta protección está directamente relacionada con la protección de la intimidad y vida privada de las personas.

Resaltando que el derecho a la privacidad surge, por regla general, del derecho a que dicha información sea individual y personal, y no debe ser pública, es derecho de todo ciudadano: saber, rectificar y determinar quién, cómo, cuándo y cómo se divulgarán sus datos personales, de conformidad con la LGPD (Brasil, 2022). Desde la llegada de Internet se resignificó el derecho a la privacidad, haciéndolo más volátil.

Es importante señalar que la necesidad de protección jurídica del ciudadano parte de la constatación de que los datos personales que circulan en la web tienen un valor económico, es decir, donde existe la posibilidad de comercializar dicha información. El impacto de la LGPD (Brasil, 2022), observando la evolución tecnológica y la necesidad de innovación de las empresas ante este nuevo escenario, orienta los efectos económicos, sociales y políticos en el país. De ahí la necesidad de profesionales con conocimientos especializados en temas específicos relacionados con la protección de datos personales, lo que en la LGPD (Brasil, 2022) se denomina DPO (Data Protection Officer).

El DPO – Oficial de Protección de Datos tendrá el desafío de coordinar y adecuar las empresas no solo a la LGPD(Brasil, 2022), sino a otras leyes que tratan el tema, como el Código de Protección al Consumidor, Marco Civil da Internet(Souza, J. S. de; et al, 2020), Constitución Federal (Brasil, 1988); Además de las leyes, también existen estándares de seguridad de la información (estándares normativos, entre otros) (Souza, J. S. de; et al, 2020) que deben orientar las prácticas de seguridad aplicadas. Para cumplir con la Ley, el DPO necesita mapear los procesos y comprender el flujo de información dentro de la organización, quiénes son las personas involucradas y estudiar las tecnologías para soportar todo el entorno.

Se promulgó una nueva ley en Brasil, la Ley General de Protección de Datos Personales, Ley N° 13.709 (Brasil, 2018), de 14 de agosto de 2018, que asegura a la población brasileña derechos y garantías sobre cómo las organizaciones en materia de recolección y tratamiento de datos datos personales, ya sea por medios físicos o digitales, así como en el Metaverso. Por lo tanto, las empresas comenzaron a preocuparse por proteger los datos de las personas debido a la aprobación de la Ley brasileña. Brasil se destaca

en América Latina en términos únicos en términos de sensibilización, regulación de nuevos negocios, ya que sus profesionales tienen formación internacional y son reconocidos oficialmente, ya que la ocupación de DPO-Oficial de Proceso de Datos está incluida en la Clasificación Brasileña de Ocupaciones (CBO). El *Data Officer* tiene un papel vital previsto en el Artigo. 41 de la LGPD (Brasil, 2018), para evitar el riesgo de altas multas impuestas por el incumplimiento de sus normas, siendo un profesional preparado para que las empresas se adapten a la Ley.

La relevancia de la Privacidad de Datos ganó más evidencia luego de los escándalos que involucraron a empresas dentro de las redes sociales acusadas de vender datos personales a empresas de marketing.

El cambio de postura es una acción mundial en materia de seguridad de la información, en todos los países se ha creado una estricta legislación sobre privacidad de datos. Por ejemplo, el RGPD, el Reglamento Europeo sobre Privacidad de Datos Personales que sirvió de modelo para la ley brasileña LGPD – Lei Geral de Privacy de Dados Personal(Brasil, 2018). Ambas leyes sugieren que la adecuación de la organización necesita involucrar áreas como TI, Seguridad de la Información y el área Jurídica; para ello, traen como requisito un profesional responsable de la privacidad de los datos personales, el DPO – Delegado de Protección de Datos.

Algunos ejecutivos académicos y entusiastas que formaron parte de los comités técnicos en Brasilia señalaron que era hora de que los profesionales de TI tengan una representación enfocada en el tema de privacidad y privacidad de datos. En Brasil, los profesionales de DPO están asociados a ANPPD - Asociación Nacional de Profesiones de Privacidad de Datos.

En las disposiciones preliminares de la LGPD (Brasil, 2018), exactamente en el art. 6 sobre actividades de tratamiento de datos personales, donde destaca que los principios de buena fe, sobre seguridad de la información deben utilizar medidas técnicas y administrativas capaces de proteger estos datos de accesos no autorizados e incidentes que produzcan la pérdida, alteración, comunicación o difusión (Brasil, 2018), precisamente el propósito de este artículo dentro del entorno del Metaverso.

También se destacan los artículos 9 y 11 de la LGPD, referidos al tratamiento de datos personales sensibles, que son aquellos que pueden discriminar a una persona, como el origen racial o étnico y la convicción religiosa, los cuales solo pueden ser manipulados en caso de prevención del fraude. , seguridad del titular de los datos, proceso de identificación y autenticación catastral en sistemas electrónicos, además de excepciones cuando exista la necesidad de garantizar la protección de los datos de dicho titular (Brasil, 2018).

Según el artículo 46 de la LGPD (Brasil, 2018), existe un enfoque de seguridad y confidencialidad de los datos, donde los agentes del proceso deben utilizar medidas de seguridad, técnicas y administrativas capaces de proteger el acceso a los datos personales no autorizados y de incidentes accidentales o ilícitos, tales como la pérdida, alteración y comunicación, o trato inadecuado (Brasil, 1988).

1.3 Inteligencia Artificial

Alan Turing realizó los primeros trabajos en el área de la Inteligencia Artificial, aunque todavía se ve como algo reciente, estos estudios comenzaron después de la Segunda Guerra Mundial. Incluso en los años 50 se desarrollaron lenguajes de programación específicos como Lisp para este fin. A partir de estos estudios se llevaron a cabo varias investigaciones. Es difícil definir el concepto de inteligencia artificial, es decir, es una noción que tiene múltiples interpretaciones, muchas veces conflictivas o circulares, porque describir qué es la inteligencia ya es un desafío, aún más complejo cuando se trata de una implementación artificial, de lo cual debe sustentarse en conceptos deterministas y empíricos.

Como las facultades humanas están absolutamente inmersas en la subjetividad, situación que dificulta reproducirlas involucrando una visión, análisis y síntesis puramente digital o computacional, a pesar del uso de redes neuronales artificiales, lógica difusa, entre otros. Básicamente, la Inteligencia Artificial, a pesar de todas estas barreras, pretende reproducir el pensamiento humano.

En el umbral de este período, la Inteligencia Artificial está enfocada a reproducir facultades humanas como la creatividad, la superación personal y el uso de lenguajes de Redes Neuronales Artificiales, como está presentando ChaptGPT, que se alimenta de la información que captura de Internet.

Al acercarse, la arquitectura de red de la Red Neuronal Recurrente (RNN), donde las neuronas de la capa oculta de la red neuronal recurrente reciben el resultado de la operación matemática que realizaron en el tiempo anterior, con los datos de la capa anterior. Por lo tanto, estas redes pueden modelar problemas con características temporales, como el pronóstico del tiempo, a partir de los resultados de períodos anteriores. Así, las RNNs consideran una dependencia temporal entre los datos de entrada.

1.4 Machine Learning

Las redes neuronales pueden "aprender" a evaluar la toma de decisiones a partir del análisis de varios resultados previos y sus consecuencias, por lo que dependen de pruebas preliminares con una base de datos estructurada de información proporcionada por expertos sobre casos reales. Existen varios métodos para entrenar una red neuronal, sin embargo, se debe tener precaución con el “sobreentrenamiento”, es decir, un entrenamiento excesivo o dirigido, que conduciría a resultados fuera de lo esperado.

La función de activación recibe la suma de la multiplicación de cada peso por el respectivo valor del parámetro (o señal) evaluado. Esto se logra punteando la matriz de peso con la matriz de valor del parámetro. El resultado de la suma sirve como parámetro de entrada para una función de activación, donde se define un umbral de activación.

Si el resultado del cálculo de la función de activación supera el umbral, se propaga un valor a la siguiente neurona por delante. La función de activación no lineal suele dar mejores resultados, como en la figura 1:

Función de activación Sigmoid σ (z): (rango de 0 a +1): notación de activación utilizada en la función lineal. Para cada z mayor, tenemos σ (z): = 1; para z, menor o tendiendo a menor, si σ (z) ≈0. En el entrenamiento, el sigmoide promedio es de alrededor de 0,5. Por lo general, solo se usan en la capa de salida cuando se espera un binario.

Tanh-función de activación hiperbólica (rango de -1 a +1): notación de activación utilizada en funciones no lineales. Por lo general, se usa en la capa de salida y oculta (aprendizaje) porque na está cerca de 0.0.

Función de activación Unidad lineal rectificada ReLU: notación de activación utilizada en la función no lineal. Usualmente se usa en la capa oculta (aprendizaje).

Función de activación (ver figura 1) ReLU-Unidad lineal rectificada con fugas [Suelta] (máx.): notación de activación utilizada en la función no lineal. Suele usarse en la capa oculta (aprendizaje), así σ (z): = 0.01.

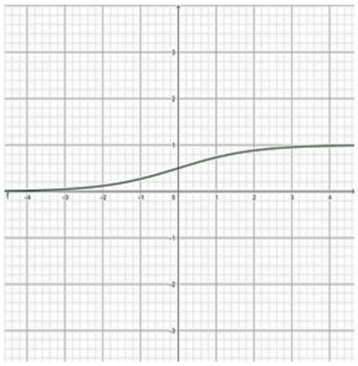


Figura 1 Sigmoid (De Lima, 2021): $a = \frac{1}{1 + e^{-z}}$

1.5 Ley General de Protección de Datos x Agencia Nacional de Protección de Datos y el Metaverso

Este artículo aborda la importancia de crear sistemas que tengan como objetivo dar la debida protección a los datos del individuo, dado que vives una cuarta revolución y dicha innovación afecta directamente al entorno virtual, con el desarrollo de la tecnología de la información y el desarrollo de la sociedad y la economía global.

Para vetar ilegalidades, surge el marco legal de internet, la Ley General de Protección de Datos(Brasil, 2018) que creó la Agencia Nacional de Protección de Datos, que es nuestra agencia reguladora en esta área, en Brasil, y su creación es un paso esencial para la regulación de la Protección de Datos, he aquí, con la creación de la ANPD, tenemos nuestro órgano supervisor.

La ANPD, que es una agencia federal de carácter especial en Brasil, vinculada al Ministerio de Justicia y Seguridad Pública, es responsable de supervisar, implementar y supervisar el cumplimiento de la LGPD (Brasil, 2018) en el país.

Por lo tanto, los sistemas de protección deben basarse en el Marco de Derechos Civiles Intert para Intert, la Ley General de Protección de Datos, las regulaciones emitidas por la ANPD, el Código Civil, la Ley del Consumidor y otras leyes.

Con respecto a los metaversos x la LGPD (Brasil, 2018), podemos verificar el cuidado con respecto al intercambio de datos que pueden identificar a un individuo, este individuo clasificado de acuerdo con la Constitución Federal y el Código Civil, comprobando así la importancia del blindaje de los datos, para cada individuo tiene el derecho de personalidad protegido en la nueva herramienta de IA, que es el metaverso.

Con respecto al individuo, el Código Civil en su artículo 2 define cuándo debe iniciarse la personalidad civil de la persona y trata sus derechos en el capítulo de la personalidad entre los artículos 11 a 21, tenemos la base de los derechos fundamentales del individuo de manera integral en la Constitución Federal, artículo 5, caput, pero usaremos aquí la idea de individuo que fue descrita por (Albuquerque & Reale, 2004), "La persona es el valor fuente de todos los valores, siendo la principal base jurídica del sistema jurídico", es decir, sin el individuo no hay sociedad.

De esta manera los derechos de la personalidad son inherentes a la persona humana, por lo que es innata.

Como se señaló anteriormente, los derechos de la personalidad son peculiares de la persona y su dignidad, por lo que surgen cinco derechos principales, a saber: el derecho a la vida, al honor, a la imagen, al nombre y a la intimidad, y estas expresiones demuestran la concepción de estos derechos.

La Ley General de Protección de Datos – LGPD (Brasil, 2018), se ocupó del blindaje del derecho del individuo al establecer en su artículo 1 que la ley tiene como objetivo proteger los derechos fundamentales de libertad y privacidad y el libre desarrollo de la personalidad de la persona física.

Crear reglas para el procesamiento de datos que eviten la fuga de la información de la persona física que viene a través de esto afectar directamente la integridad de la personalidad de la persona.

De ahí la importancia de explorar el estudio descrito en este artículo sobre la optimización de la seguridad en el metaverso utilizando DLP – Data Loss Prevention junto con la lógica paraconsistente anotada evidencial Eт.

Entonces se sigue que hoy contamos con medios legales creados para la supervisión y aplicación de las normas definidas en la Ley General de Protección de Datos y las demás normas de derecho, y a través de la Agencia Nacional de Protección de Datos, tenemos nuestro órgano supervisor que tiene el deber de multar a quienes no realicen un sistema de seguriad para la protección de datos definido en la Ley.

Con el avance de la tecnología hemos llegado al metaverso que ha traído consigo un nuevo fenómeno social de relación, donde lo real se acerca a lo virtual y viceversa, y puede afectar directamente el derecho de la

personalidad, he aquí, a acceder a datos sensibles que se comparten y si no se tratan correctamente y existiendo la imprudencia o negligencia de la persona responsable de recopilar los datos, El daño irreparable ocurrirá al individuo, por eso se presenta este artículo el estudio de la optimización de la Seguridad en el Metaverso, con el uso de sistemas que ayudan a prevenir la fuga de datos.

2. Revisión de La Literatura

2.1 Metaverso

La curiosidad de la humanidad por buscar "mundos paralelos" siempre ha sido un desafío, y con la creación del Metaverso surge la oportunidad de experimentar y convivir con un ciberespacio o entorno digital. Este tipo de experiencia conduce a una experiencia de Sociedad Red (Schlemmer, 2008).

Esta materialización de mundos paralelos en internet, el Metaverso constituye la visión del progreso de la tecnología para crear una vida social y organizacional inmersiva con el uso de la inteligencia artificial (Samarnggoon, 2023).

Son varios los modelos de Metaverse que se han desarrollado como plataformas prácticas basadas en juegos juveniles (por La Fuente Prieto, 2022), tales como: Second Life, Meta Horizons (Facebook), Fortnite, Pokemon GO, Spatial, Sandbox, Somnium Space, Decentraland , Cryptovoxels, Axie Infinity, Alien Worlds, Illuvium, Bloktopia, Star Atlas, Roblox. Por ejemplo, Somnium Space presenta billeteras únicas que compran o venden NFT (Ante, 2022).

El ambiente del Metaverso con el uso de estímulos visuales provoca un aumento de los estados emocionales que alcanzan una experiencia de realidad cercana, comprobada por el análisis de estímulos incluso con el apoyo de la Electroencefalografía, tal como lo presenta Daşdemir,2022.

Algunos Museos ya presentan el contenido de la exhibición en Metaverse, combinando realidad aumentada y un mundo virtual (Choi, 2017) Otras empresas lo están utilizando para eventos para promocionar sus propias marcas, como Gucci, que dentro de los juegos tramita la venta de productos (virtual activos con NFT) como sus bolsos incluso con un valor más caro que el producto real en sus tiendas (Forbes, 2022).

También está el planteamiento de que en el futuro los datos personales de cada individuo tendrán un NFT que los identificaría como únicos e intransferibles de esa persona, lo que evitaría la comercialización indebida de esa información, especialmente si contiene datos sensibles como datos médicos. información, permitiendo identificar también su veracidad y debido uso (Skalidis, 2022).

Este es un ejemplo de una sala de formación para LGPD en Metaverso con la participación de OPD del Comité Científico de la ANPPD, utilizando la aplicación Spatial (Figura 2 y 3):

Figura 2 reunión del comité científico de la ANPPD en el metaverso - autores

Figura 3 reunión del comité científico de la ANPPD en el metaverso - autores

En la mayoría de los estudios referentes al Metaverso, se observa que existe la necesidad de desarrollar políticas que traten el medio ambiente de manera regulatoria e incluso solo en términos de seguridad de la información (Bourlakis et al, 2009).

2.2 DLP – Data Loss Prevention

El concepto de DLP – Data Loss Prevention no es nuevo, pero se está utilizando cada vez más en la actualidad, ya que las herramientas listas para usar en el mercado ya tienen algunas funcionalidades de este estándar de monitoreo integrado. Aplicaciones como Microsoft Office 365 ya son capaces de traer algunas plantillas estandarizadas internamente para cumplir con la LGPD u otras regulaciones como GDPR (European Comission, 2016). Incluso tratando de realizar un análisis para el cumplimiento de los estándares, con la falta de profundidad en la evaluación de los criterios de seguridad, se pueden producir pérdidas de datos (Silowash & King, 2013).

El antivirus de Trend Micro cuenta con una solución llamada Apex One, que es un SaaS – Systems as a Service para seguridad de endpoints y servidores, con la funcionalidad de reportar logs en DLP – Data Loss Prevention, que se puede configurar de acuerdo a las necesidades de la empresa, por ejemplo: busque la divulgación de información sobre CPF, CNPJ, Facturas, tarjeta de

crédito, seguridad social, NFT o incluso palabras específicas etc. (Silowash & King, 2013).

Cada una de estas herramientas utiliza algoritmos de inteligencia artificial que cruzan datos y monitorean patrones preformateados por el analista de seguridad de la información, partiendo de la lógica clásica para conclusiones y presentaciones para la toma de decisiones. Principalmente, cuando se detecta una situación de pérdida, robo o fuga de datos.

El DLP debe tener una configuración asertiva para capturar resultados alineados con las necesidades de la empresa, como la cantidad de transacciones de tarjeta de crédito por cliente, cuáles son válidas o si hay bloqueos para la misma tarjeta, entre otras situaciones, para evaluar situaciones sospechosas dieciséis (Priest et al, 2018).

DLP también se puede estandarizar para evaluar malware sospechoso, ransomware y spam, utilizando listas de seguridad pública como el sitio web de Mitre Attack.

Muchas empresas han desarrollado herramientas que logran unificar alertas sobre actividades maliciosas como escaneo de IP, flujo de datos, correos electrónicos maliciosos, intentos de intrusión, monitoreo de firewall, revisión de actualizaciones de políticas de seguridad, parches de seguridad implementados, pero en este estudio el foco está específicamente en la pérdida. , robo o fuga de activos habilitados para NFT dentro del entorno de Metaverse(Priest et al, 2018).

Además, optimizar la identificación efectiva de activos redundantes, inválidos y no NFT, su identificación, si se ha producido algún cambio o transición indebida, beneficiando el monitoreo efectivo de estos valiosos activos (Silowash & King, 2013).

Las herramientas DLP requieren pruebas y calibración, situación en la que pueden ocurrir alertas con falsos positivos excesivos, pero utilizan la lógica clásica (Sikorski, M.& Honig, A., 2012).

Dado que el nivel de sofisticación alcanzado por el software malicioso requiere esfuerzos constantes para mitigar esta práctica con soluciones preventivas, como el uso de DLP para el monitoreo efectivo en conjunto con Paraconsistent Logic, para innovar y estar un paso por delante de ellos (Singh, J. & Singh, J., 2018).

2.3 NFT - Non-Fungible Tokens

El NFT – Non-fungible token es un derecho negociable en blockchain, referido a activos digitales (imágenes, música, videos, creaciones virtuales, obras de arte y parte de mundos virtuales como Metaverso), con contratos de propiedad registrados. El precio de NFT surgió a raíz de las criptomonedas, destacándose a principios de 2021 (Dowling, 2022).

Los NFT son únicos y "no fungibles, definidos como activos digitales puros que no se pueden intercambiar como iguales, a diferencia de las criptomonedas tradicionales (por ejemplo, Bitcoin o Ethereum), cuyas

monedas son todas equivalentes, indistinguibles y "fungibles". Esto les permite demostrar la autenticidad y propiedad de diferentes elementos: eventos virtuales, coleccionables digitales (por ejemplo, tarjetas coleccionables, imágenes digitales, videos, bienes raíces virtuales, nombres de dominio y sellos criptográficos), juegos y metaversos (Wang et al, 2021).

El concepto de fungibilidad se refiere a la generalidad o sustituibilidad de bienes jurídicos o disposiciones obligatorias (Lamy, 2007).

Surge una economía creativa descentralizada y abierta para garantizar la propiedad de los activos digitales basados en NFT (García, 2022).

Ha habido un crecimiento en el comercio de NFT acompañado de un gran aumento en la discusión pública y la cobertura de los medios (Dowling2, 2022).

Un ejemplo de Metaverso es Decentraland, donde un NFT se llama LAND que puede ser negociado para que un usuario construya su espacio dentro del lote del comprador en ese entorno, ya con la perspectiva de la Directiva de Contenido Digital (Digital Content Directive) (Goanta, 2020).

Otro ejemplo son los activos denominados CryptoPunks, que se crearon en 2017 alrededor de 10.000 y cuentan con personajes únicos considerados únicos y coleccionables. Los personajes individuales de CryptoPunks en abril de 2020 se comercializaron entre US$ 50 y US$ 100 cada uno, mientras que en febrero y marzo de 2021 se comercializaron entre US$ 20 000 y US$ 100 000 (Dowling, 2022).

2.4 Logic Paraconsistent Annotated Evidential Eт

Las lógicas anotadas constituyen una clase de lógicas paraconsistentes. Tales lógicas están relacionadas con ciertas redes completas, que juegan un papel esencial. Conocimiento de un experto sobre un tema analizado, las preguntas se utilizan para captar opiniones que se normalizan en lógica entre 0 y 1, como se muestra en la Figura 3. Estos valores son, respectivamente, la evidencia favorable que expresa el símbolo μ y la evidencia contraria por λ. La lógica Eт debe seguir el proceso (ver figura 4) durante la aplicación, que se puede ver en la figura:

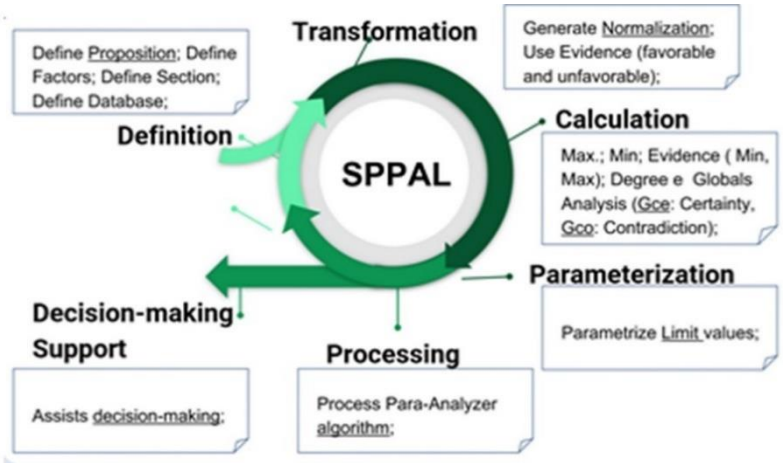


Figure 4 Passos do processo da Logica Paraconsistent Anotada Evidendial Eт

La definición significa entender y crear la proposición que se quiere entender y que debe reflejar el problema.

La transformación trataría datos tanto favorables como desfavorables. Usualmente normalizamos la respuesta entre cero y uno para ser manipulada por Eт Logic.

Por lo tanto, se hace posible procesar los cálculos de los datos recopilados. Con eso tenemos el Grado Favorable y el Grado Desfavorable.

Se parametrizan límites aceptables para obtener un análisis que haga los datos y pueda transformarse en información útil.

La inteligencia se obtiene aplicando el algoritmo para-parser que contiene toda la información para ejecutar la Eт Logic.

Como se resume en Abe et al. (Abe et al 2011), los programas ahora se pueden construir utilizando lógica paraconsistente, lo que hace posible manejar las inconsistencias de una manera sencilla y elegante. Este recurso puede aplicarse en sistemas expertos, bases de datos orientadas a objetos, representación de conocimientos contradictorios, etc., con todas las implicaciones en inteligencia artificial.

En Abe et al (2011): “"La Lógica Paraconsistente Anotada Probatoria Eт tiene un lenguaje Eт y las proposiciones atómicas son del tipo p (μ, λ) donde p es una proposición y μ, λ∈ [0, 1]. Intuitivamente, μ indica el grado de evidencia desfavorable de p y λ el grado de evidencia contraria de p.Sobre los valores μ, λ depende de las aplicaciones consideradas y puede cambiar: de hecho μ puede ser el grado de creencia favorable y λ puede ser el grado de creencia contrario a la proposición p; Además, μ puede indicar la probabilidad expresada de que ocurra p y λ la improbabilidad expresada de que ocurra p. Las proposiciones atómicas p (μ, λ) de la lógica Eт pueden leerse intuitivamente como: Me quedo en p con el grado de creencia favorable μ y el grado de creencia contraria λ, o el grado de evidencia favorable de p es μ y el grado de evidencia contraria de p es λ".

Las lógicas paraconsistentes pueden servir como la lógica subyacente de teorías en las que A y ¬A (la negación de A) son verdaderas sin ser triviales (Abe, J. M. & Nakamatsu K., 2015). Hay muchos tipos de sistemas paraconsistentes. En este texto, consideramos la Lógica de Evidencia Paraconsistente Anotada Eт. Las formulaciones en Eт Lógica son del tipo p (μ, λ), donde p es una proposición y e (μ, λ) ϵ [0, 1] es el verdadero intervalo cerrado unitario.

Una proposición p(μ, λ) puede leerse como: "La evidencia favorable de p es μ y la evidencia desfavorable es λ" (Abe et al 2011). Por ejemplo, p (1.0, 0.0) se puede leer como una proposición verdadera, p (0.0, 1.0) como falsa, p (1.0, 1.0) como inconsistente, p (0.0, 0.0) como paracompleta y p (0.5, 0.5) como una proposición indefinida (Abe et al 2011).

También introducimos los siguientes conceptos: Grado de incertidumbre: Gun (μ, λ) = μ + λ - 1 (0 ≤ μ, λ ≤ 1) y Grado de certeza: Gce(μ, λ) = μ - λ (0 ≤ μ, λ ≤ 1) (De Carvalho, F.R., et al, 2011). Una relación de orden se define en [0, 1]: (μ1, λ1) ≤ (μ2, λ2) ↔ μ1 ≤ μ2 y λ 2 ≤ λ1, constituyendo una red que será simbolizada por τ.

Con los grados de incertidumbre y certeza, podemos obtener los siguientes 12 estados de salida (Tabla 2): estados extremos y no extremos. Cabe señalar que esta división puede modificarse según cada aplicación (Subrahmanian, V, 1987).

Algoritmo del paraanalizador. En esta propuesta de algoritmo se obtiene un conjunto de información que en ocasiones puede parecer contradictoria, dificultando el análisis del escenario para el análisis de riesgo. Generalmente, en tales situaciones, esta información es descartada o ignorada, es decir, se considera "suciedad" del sistema, sin embargo, en el mejor de los casos, puede incluso recibir un tratamiento diferente.

Silva Filho, Abe y Torres (2011) citan: "Sin embargo, la contradicción, la mayoría de las veces, contiene información decisiva, ya que es como el encuentro de dos hilos de valores de verdad opuestos. anacrónico, y es por eso que debemos buscar lenguajes que puedan convivir con la contradicción sin perturbar a la otra información. En cuanto a la incertidumbre, hay que pensar en un lenguaje que pueda captar el 'máximo' de 'información' del concepto".

En esta línea de razonamiento para el análisis basado en la Lógica Paraconsistente, se considerarán situaciones de Inconsistencia y Paracompletitud junto con Verdadero y Falso, representadas según la Tabla 1:

Tabla 1 – Estados extremos Fuente: Abe et al (2011).

Estados Extremos	**Símbolo**
Verdadero	V
Falso	F
Inconsistente	T
Paracompleto	⊥

El conjunto de estos estados u objetos (τ = {F, V, T,⊥}) también se puede llamar constantes de anotación y se puede representar usando el diagrama de Hasse como se muestra en la figura 5:

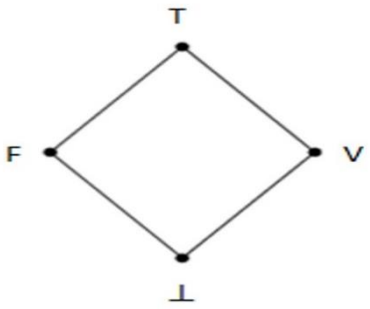


Figura 5 Diagrama de Hasse (Silva Filho, Abe y Torres, 2011)

"El operador sobre τ es: $\sim:|\tau| \rightarrow |\tau|$ que funcionará, intuitivamente, así:

$\sim$ T = T (la "negación" de una proposición inconsistente es inconsistente)

$\sim$ V = F (la "negación" de una proposición verdadera es falsa)

$\sim$ F = V (la "negación" de una proposición falsa es verdadera)

$\sim\perp = \perp$ (la 'negación' de una proposición paracompleta es paracompleta)

Se utilizará Lógica Paraconsistente Anotada; Este tipo debe estar compuesto por 1, 2 o "n" valores.

Con los cálculos de los valores de los ejes que componen la figura representativa de la red, se puede dividir o delimitar internamente en varias regiones de diferentes tamaños y formatos, obteniendo así una discretización de la misma. A partir de las regiones delimitadas de la red, es posible relacionar los estados lógicos resultantes, los cuales, a su vez, se obtendrán interpolando los Grados de Certeza Gc y Gct de Contradicción. Así, para cada interpolación entre los grados de certeza y contradicción, es posible extraer información para ayudar en la toma de decisiones (Subrahmanian, V., 1987).

La representación de la tabla 2 muestra una representación de la red construida con valores de Grados de Certeza y Contradicción y seccionada en 12 estados. Así, al final del análisis se obtendrá como respuesta para la toma de decisiones uno de los 12 posibles estados lógicos resultantes.

Tabla 2. Estados no extremos

Estados no extremos	Simbolo
Cuasi verdadero que tiende a Inconsistente	QV⟶T
Cuasi-verdadero con tendencia a Paracompleto	QV⟶⊥
Cuasi-falso tendiendo a Inconsistente	QF⟶T
Cuasi-falso tendiendo a Paracompleto	QF⟶⊥
Cuasi-inconsistente que tiende a Verdadero	QT⟶V
Cuasi-inconsistente que tiende a Falso	QT⟶ F

Cuasi-paracompleto tendiendo a Verdadero	Q⊥⟶V
Cuasi-paracompleto que tiende a Falso	Q⊥⟶ F

Algunos valores de control adicionales son:

• Vscct = valor máximo de control de incertidumbre = Ftun

• Vscc = valor máximo de certeza de control = Ftce

• Vicct = valor mínimo de control de incertidumbre = -Ftun

• Vicc = valor mínimo de certeza de control = -Ftce

Todos los estados están representados en la Figura 6 a continuación.

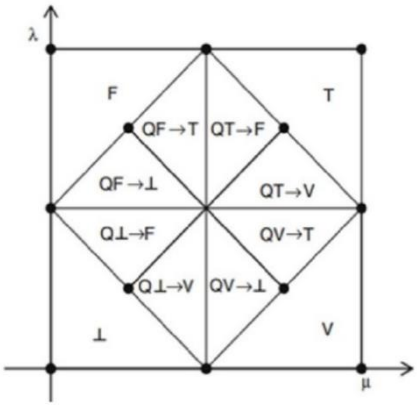


Figura 6 Aspecto de la red para tomar decisiones Fuente: (Abe et al., 2011).

2.5 Técnicas de Inteligencia Artificial

En la década de 1940, Warren McCulloch y Walter Pitts propusieron por primera vez un modelo de inteligencia artificial, como se muestra en la Fig. x, que utilizaba una estructura neuronal, sugiriendo el uso de hardware; En este caso, se comportaron resistencias variables conectadas a amplificadores (ver Figura 5). como una neurona humana (De Lima L.A., et al, 2019).

Su concepto era simple, ya que era capaz de modelar sistemas lineales separables, como los operadores lógicos AND, OR y NOT. La neurona ingresaría una lista de datos booleanos (0 o 1), haría la suma y luego, en secuencia, haría la suma a una función de activación que devolvería 1 si la suma excede el umbral y 0 si falla (De Lima L.A., et al, 2019).

Higo. X ejemplifica cómo se configuraría una neurona para calcular x1 E ! X2do. Haciendo de !x2 una entrada inhibitoria, donde solo habría dos situaciones posibles: x1 = 0 y x2 = 0 y x1 = 1 y x2 = 0. Obviamente, la expresión solo puede evaluarse como verdadera si x1 = 1 y por lo tanto el caso 2 es el único válido (De Lima L.A., et al, 2019).

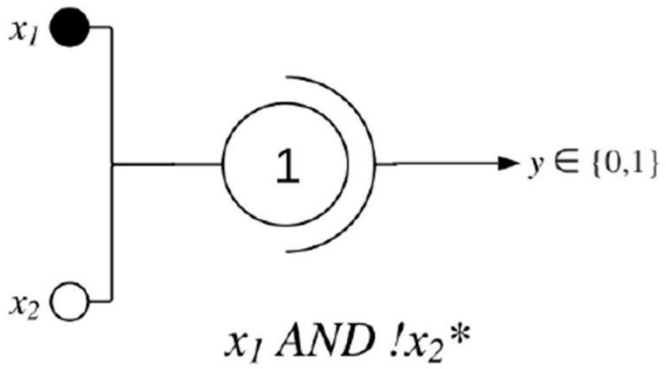


Figura 5. Modelo de neuronas de McCulloch-Pitts

Hoy en día, muchas organizaciones buscan soluciones de IA - Inteligencia Artificial para encontrar analogías que conduzcan a una contención optimizada, pudiendo utilizar el enfoque en la lógica no clásica para sortear este tipo de ataques grupales más ofensivos (Subrahmanian, V., 1987).

El uso de la Inteligencia Artificial presenta aspectos relevantes para varias áreas de investigación, incluso en el aspecto regulatorio y judicial, como algunos estudios realizados en el Supremo Tribunal Federal de Brasil (Peixoto, 2020 y Siqueira et al, 2020).

Los primeros resultados de la implementación de métodos de redes neuronales convolucionales se realizaron con el sector financiero, que aún prefiere no contratar analistas de datos exclusivos debido a este tipo de intento de ataque que utiliza Inteligencia Artificial/Redes Neuronales, a pesar de tener problemas complejos como optimización de su estructura cibernética, según un estudio presentado por la empresa Trend Micro. Incluso los resultados de las primeras pruebas demuestran el potencial de tales enfoques, que pueden estimular a las empresas a pensar en crear una IA – Inteligencia Artificial separada de la estructura interna o desarrollar una estructura de Machine Learning específica para cada cliente (De Lima L.A., et al, 2019).

La técnica de inteligencia artificial (European Comission, 2016) proviene de la década de 1940 en los estudios propuestos por Warren McCulloch y Walter Pitts (1943). Algunas investigaciones se basan en la filosofía, el conocimiento y la función de las neuronas y utilizan la lógica proposicional creada por Russell y Whitehead. Otro pilar es la teoría de la computación de Turing.

Estos dos investigadores propusieron un modelo de neuronas artificiales, en el que cada neurona se caracteriza por estar "encendida" o "apagada", ocurriendo el cambio a "encendido" en respuesta a la salida de una neurona de la capa que estimula el aprendizaje de la siguiente capa capa.

El estado de una neurona se consideró "equivalente en términos concretos a una propuesta que define su estímulo adecuado". McCulloch y Pitts también sugirieron que las redes correctamente definidas podrían aprender. Por ejemplo, demostraron que una red determinada de neuronas conectadas podía calcular cualquier función computable y que las estructuras de red simples podían implementar todos los conectores lógicos (Y, O, NO, etc.).

En la figura 6, se propone la neurona paraconsistente (Abe et al, 2011) para servir a la red neuronal paraconsistente. Otro punto importante a tener en cuenta es que la neurona elige qué característica (x) utilizará en la red a entrenar.

Figura 7 Propuesta de neurona artificial paraconsistente, (Lima, L.A. et al, 2021)

El concepto de Lógica Aplicada (De Lima, Luiz A., et al 2018) en nuestra realidad cotidiana frente a innumerables fuentes de información, la contradicción ocupa constantemente un espacio, trayendo incertidumbres que culminarán en desafíos breves o futuros.

En el caso de un sistema con inteligencia artificial (De Lima, A.W.B., et al, 2018), redes neuronales (De Lima, A.W.B., et al, 2018), también conocido como "aprendizaje automático" (Lima, Luiz A. de, et al., 2021), que parte del estudio del reconocimiento de patrones (Subrahmanian, V, 1987) (Lima et al,2020), la aparición de contradicción en la lógica del razonamiento es inevitable cuando tratamos de reflejar el comportamiento humano.

En las actividades del segmento médico, hospitalario y de salud, se ha observado en el uso de análisis de exámenes clínicos, diagnóstico precoz de cáncer " (Lima, Luiz A. de, et al., 2021), en la política, en el análisis de procesos judiciales y productividad de la seguridad pública (Insua, H. G., et al, 2022), en la medición de software (Lima, L. A., 2019) (Da Silva, J. P., 2018) de soporte técnico, al servicio de las aseguradoras, donde intervienen al menos dos especialistas (De Lima, A.W.B., et al, 2018), siempre habrá puntos de vista diferentes.

En respuesta a la contradicción, tenemos Eᴛ Logic trabajando en cualquier segmento comercial y científico que pueda agregarse con otras tecnologías. Se exploró una aplicación en servicios seis sigma, tanto en la industria como en los servicios, en la que participen al menos dos especialistas (De Lima, A.W.B., et al, 2018).

El Eт Logic ha sido investigado en el segmento de bienestar animal; siempre se ha profundizado con la previsión y toda la cadena agroindustrial (De Alencar Nääs, I., et al, 2020).

3. Optimización del monitoreo de seguridad

3.1 DLP: Prevención de Pérdida de Datos Mediante Lógica Paraconsistente Evidencial Anotada Eт

DLP: Data Loss Prevention ayuda a delimitar los activos:

• ¿Cuál es su origen?

• ¿Pasa por consistencia y validación NFT?

• ¿Cómo se hace la transición de salida de activos con NFT?

• Detección de activos con NFT no válidos.

• Pista de auditoría y registro de transacciones.

Nuestro estudio delimita los activos con NFT, ubicados dentro del Metaverso:

• En la Red Red (en los servidores de aplicaciones o bases de datos).

• En el front-end (en el servidor web).

Mientras que aquellos sin NFT o con NFT inválido son:

• ¿Están incompletos los datos de NFT?

• ¿No se validó legalmente la compra del activo?

• ¿El activo es falso?

• ¿Todavía no se ha completado la transición?

Al utilizar el concepto de DLP – Data Loss Prevention (ver figura 9), se consideró la siguiente granularidad para obtener la graduación de los tipos de pérdida de datos detectados (Bioni, B. R., 2019)

Primer nivel:

• Identificación de activos con NFT válidos, no válidos o no.

El segundo nivel (Abe, J. M. & Nakamatsu K., 2015):

• Identificación de la ubicación del activo dentro del Metaverso.

En el tercer nivel (De Lima L.A., et al, 2019):

• Validación de la masa de datos recolectados de la empresa de transporte por un solo período para ser analizados por el Algoritmo de Python para verificar activos válidos y contradictorios.

En el cuarto nivel (Subrahmanian, V., 1987), se alinea la información de excepción presentada por el informe del algoritmo de Python. Se insertan en el analizador del Algoritmo Para-Analyzer para identificar la mayor asertividad en la toma de decisiones.

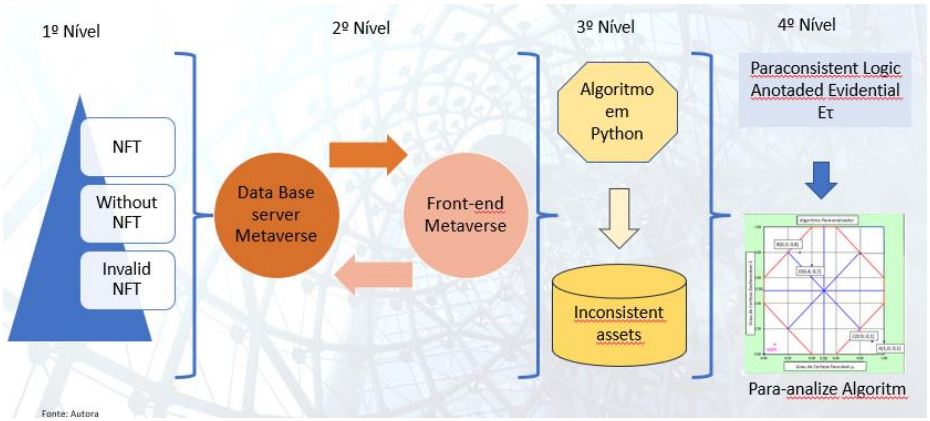


Figura 9. DLP: prevención de pérdida de datos con lógica paraconsistente anotada de evidencia Eт

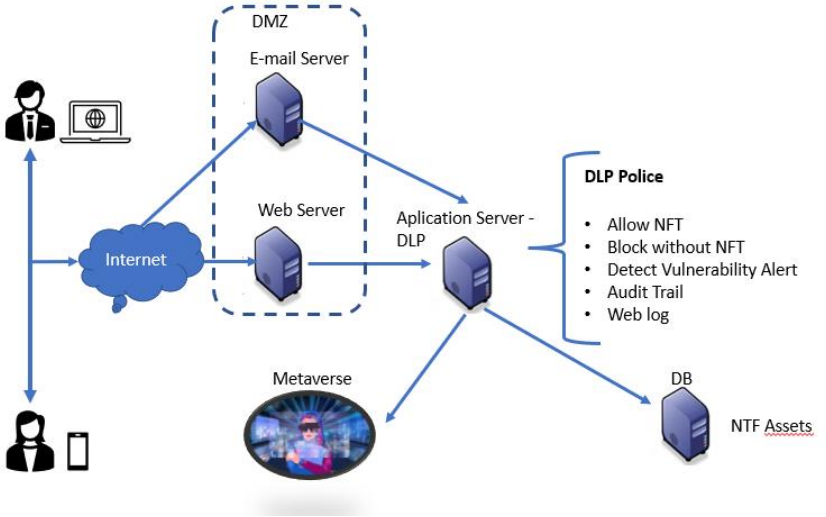


Figura 10. Estructura de la DLP – Data Loss Prevention

Las herramientas DLP - en el mercado han sido cada vez más buscadas por la empresa debido a LGPD en Brasil y GDPR en la Unión Europea-UE. Tales funcionalidades se pueden utilizar como entrada en Logic Eт.

El objetivo principal es bloquear, monitorear y gestionar permanentemente estas desviaciones.

Hay varios canales que se pueden usar que pasan por la transacción del activo, siendo plausibles para monitorear y bloquear las transferencias de NFT. Todos los detalles se realizan mediante la inspección del contenido y el contexto.

Los escaneos manuales o automáticos se realizan constantemente para excluir datos confidenciales y garantizar DLP. Hay casos en los que los activos en tránsito pueden cifrarse a la fuerza y, por lo tanto, mantener la calidad de la técnica DLP en el procesamiento de activos.

De acuerdo con la Ley, se entiende que el tratamiento debe ser considerado en todas las operaciones que se realicen con datos personales provenientes de transacciones con activos NFT, tales como las referidas a la recolección, producción, recepción, clasificación, uso, acceso, reproducción, transmisión, distribuir, procesar, archivar, almacenar, suprimir, evaluar o controlar información, modificar, comunicar, transferir, difundir o extraer.

Datos internacionales de fuera del territorio nacional y que no sean objeto de comunicación, uso compartido de datos con agentes de procesamiento brasileños u objeto de transferencia internacional de datos con otro país que no sea el país de origen, siempre que el país de origen proporcione grado de protección de datos personales adecuada a lo dispuesto en la Ley.

El uso de herramientas DLP se enfoca en cumplir con la normativa en materia de datos personales sensibles: datos personales sobre origen racial o étnico, convicción religiosa, opinión política, afiliación sindical u organización de carácter religioso, filosófico o político, datos referentes a la salud o vida sexual, datos genéticos o biométricos cuando se vinculan a un individuo.

Las empresas deben utilizar el concepto de DLP, ya que tratan los datos de acuerdo con los preceptos del operador y del controlador. Se considera responsable de las decisiones relativas al tratamiento de datos personales a una persona física o jurídica, pública o privada. Y cuando el operador sea una persona física o jurídica, pública o privada, que trate datos personales por cuenta del responsable del tratamiento.

3. Pruebas

3.1 DLP: Programa Python y resultado de datos masivos

La masa de datos a probar tenía el siguiente diseño con esta leyenda:

El primer nivel (tipo de activo):

• Activos con NFT = 1.1

• Activos sin NFT =1.2

• Activos con NFT no válido = 1,3

El segundo nivel (ubicación):

• Dentro de la base de datos del entorno del metaverso = 2.1

• Dentro de la interfaz de Metaverse = 2.2

El tercer nivel (validación):

• activos válidos =3.1

• activos contradictorios=3.2

El cuarto nivel (lógica paraconsistente):

• Verdadero=4.1

• Falso=4.2

• Incompleto=4.3

• Paracompleto=4.4

Para lo cual cada (ver figura 10) de las líneas capturadas presentó una descripción del origen de los activos.

Level 1	Level 2	Level 3	Level 4
1.1	2.1	3.1	4.1.
1.2	2.2	3.2	4.2
1.3	2.1	3.2	4.4
1.1	2.1	3.2	4.1
1.2	2.1	3.2	4.3
1.3	2.1	3.2	4.2
1.1	2.1	3.2	4.1
1.2	2.1	3.1	4.3
1.3	2.1	3.1	4.2
1.1	2.1	3.1	4.1
1.2	2.1	3.1	4.3
1.3	2.1	3.1	4.2
1.1	2.1	3.1	4.1
1.2	2.1	3.1	4.3
1.3	2.1	3.1	4.4
1.1	2.2	3.1	4.1
1.2	2.2	3.1	4.3
1.3	2.2	3.1	4.4
1.1	2.2	3.1	4.1
1.2	2.2	3.1	4.3

Figura 10. Ejemplo de datos masivos

La masa de datos capturados presentó alrededor de 30 datos por día durante un mes, comprendiendo una cantidad de 200 datos para análisis, de los cuales el algoritmo verificó el 63% en Python (ver figura 11) como efectivos y el 37% como contradictorios.

Los datos considerados contradictorios fueron pasados al algoritmo del paraanalizador, obteniendo un diferencial de 9% verdadero, 6% incompleto, 10% paracompleto y 15% falso.

Sin embargo, el intervalo considerado falso puede ser objeto de un cribado más exigente que requiera un seguimiento más eficaz por el porcentaje de ocurrencias que podría ser una alerta de fuga de datos imperceptible, pero que se produce de forma sutil.

```python
import csv
from typing import List
from util import normalize_by_feature_scaling
from network import Network
from random import shuffle

if __name__ == "__main__":
    nft_parameters: List[List[float]] = []
    nft_classifications: List[List[float]] = []
    nft_species: List[int] = []
    with open(nft.csv, mode='r') as nft_file:
        nft: List = list(csv.reader(nft_file, quoting=csv.QUOTE_NONNUMERIC))
        shuffle(nft) # get our lines of data in random order
        for nft in nfts:
            parameters: List[float] = [float(n) for n in nft[1:14]]
            nft_parameters.append(parameters)
            species: int = int(nft[0])
            if species == 1:
                nft_classifications.append([1.0, 0.0, 0.0])
            elif species == 2:
                nft_classifications.append([0.0, 1.0, 0.0])
            else:
                nft_classifications.append([0.0, 0.0, 1.0])
            nft_species.append(species)
    normalize_by_feature_scaling(nft_parameters)
# Lista [camada de entrada,camada intermediaria, camada de saida], taxa de Aprendizagem
    nft_network: Network = Network([13, 7, 3], 0.9)

    def nft_interpret_output(output: List[float]) -> int:
        if max(output) == output[0]:
            return 1
        elif max(output) == output[1]:
            return 2
        else:
            return 3


    # train over the first 150 nft 10 times
    nft_trainers: List[List[float]] = nft_parameters[0:150]
    nft_trainers_corrects: List[List[float]] = nft_classifications[0:150]
    for _ in range(10):
        nft_network.train(wine_trainers, nft_trainers_corrects)

    # test over the last 28 of the nft in the data set
    nft_testers: List[List[float]] = nft_parameters[150:178]
    nft_testers_corrects: List[int] = nft_species[150:178]
    nft_results = nft_network.validate(nft_testers, nft_testers_corrects, nft_interpret_output)
    print(f"{nft_results[0]} correct of {nft_results[1]} = {nft_results[2] * 100}%")
```

Figura 12. Programa Python

4. Conclusión

Se concluye que el alineamiento entre el uso de DLP - Data Loss Prevention con la Lógica Paraconsistente presenta una ganancia significativa para el tema de monitoreo y análisis de la pérdida de activos NFT en el ambiente Metaverso.

La implementación de este tipo de herramientas puede aumentar el desempeño de Metaverso en relación al nivel de seguridad de la información, que despreciaba el 40% de los datos como intrascendentes a solo el 15%, considerando que incluso estos intervalos pueden ser monitoreados para que su efectividad sea del 100%. % de los datos transitados, mantenidos almacenados y enviados por correo electrónico o redes sociales analizados.

Se entiende que la combinación de profesionales de la privacidad de datos con la mejora de herramientas como DLP con el uso de Paraconsistent Logic, trae la posibilidad de avances significativos en la adecuación de la LGPD y un aumento en el flujo de nuevos negocios en el entorno de el Metaverso de forma segura.

Referências bibliográficas

Abe, J.M., et al. (2011) Lógica Paraconsistente Anotada Evidencial Et, pp. 38–39. Comunnicar, Santos, 2011.

Abe, J. M. Nakamatsu K. (2015) Introduction to Annotated Logics - Foundations for Paracomplete and Paraconsistent Reasoning, Series Title

Intelligent Systems Reference Library, Volume 88, Publisher Springer International Publishing, Copyright Holder Springer International Publishing Switzerland, eBook ISBN 978-3-319-17912-4, DOI 10.1007/978-3-319-17912-4, Hardcover ISBN 978-3-319-17911-7, Series ISSN 1868-4394, Edition Number 1, 190 pages, 2015.

Albuquerque, R. D. A. C. D., & Reale Júnior, M. (2004). A criminalidade informática.

Akama, S. (2016) Towards Paraconsistent Engineering, Intelligent Systems Reference Library, Germany: Springer (2016).

ALENCAR, Roldan. (2023) Direitos da Personalidade, disponível em http://www.ebah.com.br/content/ABAAABDpwAB/direitos-personalidade. Accessed on: 28 de maio 2023.

Ante, L. (2022). Non-fungible token (NFT) markets on the Ethereum blockchain: Temporal development, cointegration and interrelations. Economics of Innovation and New Technology, 1-19.

Bourlakis, M., Papagiannidis, S., & Li, F. (2009). Retail spatial evolution: paving the way from traditional to metaverse retailing. Electronic Commerce Research, 9, 135-148.

Brasil. (2018) Lei Geral de Proteção de Dados Pessoais (LGPD). Lei nº 13.709, de 14 de agosto de 2018. Available in:http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm . Accessed on: 21/04/2022.

Brasil. (1988) Constituição Federal, de 1988. Available in: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Accessed on: 09/02/2023.

Choi, H. S., & Kim, S. H. (2017). A content service deployment plan for metaverse museum exhibitions—Centering on the combination of beacons and HMDs. International Journal of Information Management, 37(1), 1519-1527.

Daşdemir, Y. (2022). Cognitive investigation on the effect of augmented reality-based reading on emotion classification performance: A new dataset. Biomedical Signal Processing and Control, 78, 103942.

Da Silva, Jonas P., Jair M. Abe, Luiz A. De Lima, Felipe S. David De Oliveira, Kazumi Nakamatsu. (2018) "Use of Software Metrics to Scope Control in IT Projects Using Paraconsistent Logic". Journal WSEAS Transactions on Computer Research. WSEAS Transactions on Computer Research, ISSN/E-ISSN:1991-8755/2415-1521, Volume 6, 2018, Art. #8, pp. 55-59. (2018). https://www.wseas.org/multimedia/journals/computerresearch/2018/a145918-057.php

De Alencar Nääs, Irenilza; et al. (2020) "Lameness prediction in broiler chicken using a machine learning technique". INFORMATION PROCESSING IN AGRICULTURE, v. 1, p. '-13, 2020. DOI: /doi.org/10.1016/j.inpa.2020.10.003 https://linkinghub.elsevier.com/retrieve/pii/S2214317320302092

De Carvalho, F.R., Abe, J.M. (2011) Tomadas de decisão com ferramentas da Lógica Paraconsistente Anotada. São Paulo. Blucher, pp. 37–47, 2011.

De Carvalho, F.R., Brunstein, I., Abe, J. M. (2011) Paraconsistent Annotated Logic in Analysis of Viability: in approach to product launching. In: Dubois, D.M. (ed.), vol. 718, pp. 282–291, 2011.

De Lima L.A., Abe J.M., Martinez A.A.G., de Frederico A.C., Nakamatsu K., Santos J. (2019) "Process and Subprocess Studies to Implement the Paraconsistent Artificial Neural Networks for Decision-Making". In: Jain V., Patnaik S., Popențiu Vlădicescu F., Sethi I. (Eds) Recent Trends in Intelligent Computing, Communication and Devices. Advances in Intelligent Systems nd Computing, Vol 1006. Springer, Singapore. 2019 Print ISBN: 978-981- 13-9405-8; Online Isbn: 978-981-13-9406-5; HTTPS://DOI.ORG/10.1007/978-981-13-9406-5_61

De Lima, Luiz A.; Abe, Jair M.; Kirilo, Caique Z.; Da Silva, Jonas P.; Nakamatsu, Kazumi. (2018) "Using Logic Concepts in Software Measurement." PROCEDIA COMPUTER SCIENCE, v. 131, p. 600-607, 2018. http://dx.doi.org/10.1016/j.procs.2018.04.302

De Lima, A.W.B., de Lima, L.A., Abe, J.M., Gonçalves, R.F., Alves, D. and Nakamatsu, K. (2018) "Paraconsistent Annotated Logic Artificial Intelligence Study In Support Of Manager Decision-Making". IN: THE 2ND INTERNATIONAL CONFERENCE, 2018, BARCELONA. PROCEEDINGS OF THE 2ND INTERNATIONAL CONFERENCE ON BUSINESS AND INFORMATION MANAGEMENT - ICBIM' 18. BARCELONA, SPAIN: ACM DL, 2018. P. 154-157. DOI:dx.doi.org/10.1145/3278252.3278269. https://dl.acm.org/doi/10.1145/3278252.3278269

Dowling, M. (2022). Is non-fungible token pricing driven by cryptocurrencies?. Finance Research Letters, 44, 102097.

Dowling, M. (2022). Fertile LAND: Pricing non-fungible tokens. Finance Research Letters, 44, 10296.

European Comission Guidelines on Data Protection Officers ('DPOs') (wp243rev.01) 2016 Available in :<https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612048> Accessed on: 21/04/2023.

European Comission GDPR - General Data Protection Regulation 2016 Available in :< https://gdpr-info.eu/ > Accessed on: 21/04/2022.

Forbes. (2022)Available in: https://forbes.com.br/forbes-tech/2022/01/exemplos-do-metaverso-marcas-que-atuam-com-propriedade/. Accessed on: 20/05/2023.

García, R., Cediel, A., Teixidó, M., & Gil, R. (2022). Semantics and non-fungible tokens for copyright management on the metaverse and beyond. arXiv preprint arXiv:2208.14174

Goanta, C. (2020). Selling LAND in Decentraland: The regime of non-fungible tokens on the Ethereum blockchain under the digital content directive. Disruptive technology, legal innovation, and the future of real estate, 139-154.

Insua, Hugo Gava, Jair M. Abe and Luiz A.de Lima. (2022) "Produtividade da Polícia Civil do Estado de São Paulo: uma Análise". IJDR-International Journal of Development Research. ISSN: 2230-9926, Volume:12, Article ID:23962, 4 pages, Research Article. 2022. HTTPS://DOI.ORG/10.37118/IJDR.23962.02.2022

Lamy, Eduardo de Avelar. (2007) Princípio da Fungibilidade no Processo Civil. São Paulo (SP): Dialética, 2007

Lima, L. A. de; et al. (2020) "DPO no Brasil sob a ótica da LGPD - Lei Geral de Proteção de Dados". Instituto EXIN - Ministry of Economic Affairs in the

Netherlands. 2020. https://www.exin.com/br-pt/dpo-no-brasil-sob-a-otica-da-lgpd-lei-de-protecao-de-dados/

Lima, L. A., Abe, J. M., Martinez, a. A. G., Santos, J., Albertini, G., & Nakamatsu, K. (2019). "The Productivity Gains Achieved in Applicability of The Prototype AITOD with Paraconsistent Logic in Support in Decision-Making in Project Remeasurement". Proceedings of the 9th International Conference of Information and Communication Technology [ICICT-2019] Nanning, Guangxi, China January 11-13, 2019 (http://aivr.org/index.html). Edited by Srikanta Patnaik Volume 154, Pages 1-844 (2019). Procedia Computer Science, pp. 347–353. https://doi.org/10.1016/j.procs.2019.06.050

Lima, Luiz A. de, et al. "Application of architecture using AI in the training of a set of pixels of the image at aid decision-making diagnostic câncer". 25th International Conference on Knowledge Based and Intelligent Information and Engineering Systems (KES2021). 8th – 10th September 2021 | Szczecin, Poland & Virtual. IS27: Reasoning-based Intelligent Applied Systems: HTTP://KES2021.KESINTERNATIONAL.ORG/CMSISDISPLAY.PHP

Lima, Luiz A., et al. (2021) "Study of PANN Components in Image Treatment for Medical Diagnostic Decision-Making". N.70. The 2nd International Conference on Network Enterprises & Logistics Management - NETLOG 2021. ISSN 2595-0738. HTTP://WWW.NETLOGCONFERENCE.COM/PAPERS.HTML

Peixoto, F. H. (2020). Projeto Victor: relato do desenvolvimento da inteligência artificial na repercussão geral do Supremo Tribunal Federal. Revista Brasileira de Inteligência Artificial e Direito-RBIAD, 1(1), 1-22.

Priest, Graham, Koji Tanaka, and Zach Weber, (2018) "Paraconsistent Logic", The Stanford Encyclopedia of Philosophy (Summer 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2018/entries/logic-paraconsistent/>. ISSN: 1095-5054

Samarnggoon K, Grudpan S, Wongta N and Klaynak K. (2023). Developing a Virtual World for an Open-House Event: A Metaverse Approach.Future Internet. 10.3390/fi15040124. 15:4. (124).

Schlemmer, E., & Backes, L. (2008). Metaversos: novos espaços para construção do conhecimento. Revista Diálogo Educacional, 8(24), 519-532.

Silowash, George J.; King, Christopher. (2013) Insider threat control: Understanding data loss prevention (DLP) and detection by correlating events from multiple sources. CARNEGIE-MELLON UNIV PITTSBURGH PA SOFTWARE ENGINEERING INST, 2013.

Sikorski, M., Honig, A. (2012) "Practical Malware Analysis: The Hands-On Guide to Dissecting Malicious Software", 2012, No Starch Press.

Singh, J., Singh, J., (2018)"Challenges of Malware Analysis: Obfuscation Techniques", 2018. Disponível em: http://50.87.218.19/ijiss/index.php/ijiss/article/view/327

Siqueira, D. P., Lara, F. C. P., & Lima, H. F. C. (2020). Acesso à justiça e inteligência artificial: abordagem a partir da revisão sistemática da literatura. Revista Argumentum-Argumentum Journal of Law, 21(3), 1265-1277.

Skalidis, I., Muller, O., & Fournier, S. (2022). The metaverse in cardiovascular medicine: applications, challenges, and the role of non- fungible tokens. Can J Cardiol, 38(9), 1467-8.

Souza, J. S. de; et al.(2020) "The General Law Principles for Protection the Personal Data and their Importance". In: 7th International Conference on Computer Science Engineering and Information Technology (CSEIT 2020 - https://arxiv.org/abs/2009.14313), 2020. Computer science & information technology (cs & it), Copenhagen, Denmark. Anais 2020. V. 10. P. 109. HTTP://DX.DOI.ORG/10.5121/CSIT.2020.101110

Subrahmanian, V. (1987) On the semantics of quantitative logic programs. In: Proceedings of the 4th IEEE Symposium on Logic Programming, pp. 173–182 (1987)

Wang, Q., Li, R., Wang, Q., & Chen, S. (2021). Non-fungible token (NFT): Overview, evaluation, opportunities and challenges. arXiv preprint arXiv:2105.07447.