FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

Assignment of master's thesis

Title:	Machine learning based approach for summarizing governance proposals for decentralized autonomous organizations
Student:	Bc. Herman Tiumentsev
Supervisor:	Mgr. Alexander Kovalenko, Ph.D.
Study program:	Informatics
Branch / specialization:	Knowledge Engineering
Department:	Department of Applied Mathematics
Validity:	until the end of summer semester 2023/2024

Instructions

The thesis aims to develop a personalized system that summarizes proposals related to decentralized autonomous organizations (DAO) governance, with the goal of increasing accessibility and participation in the decision-making process. The research will explore different summarization approaches, including abstractive and extractive summarization, and determine the most effective machine learning-based method for summarizing complex documents related to DAO governance. As a thesis byproduct, a dataset containing text summarization will be produced.
- Conduct a thorough review of the existing literature on DAO, their governance structures, and the decision-making processes involved.
- Identify the current challenges and limitations related to summarizing proposals related to DAO governance, and how they impact accessibility and participation in the decision-making process.
- Evaluate different summarization approaches, such as abstractive and extractive summarization, and determine their advantages and disadvantages in the context of DAO governance proposals.
- Determine the criteria for evaluating the effectiveness of a summarization system for DAO governance proposals, such as accuracy, comprehensibility, and relevance.
- Develop a customized system for summarizing DAO governance proposals that address the limitations of existing approaches and meet the identified criteria for effectiveness.

Electronically approved by Ing. Magda Friedjungová, Ph.D. on 21 February 2023 in Prague.

FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

- Test the system on a sample of DAO governance proposals, and evaluate its effectiveness using the identified criteria.
- Compare the results of the customized machine learning-based system with existing approaches to summarize DAO governance proposals to demonstrate the improvements in accessibility and participation in the decision-making process.
- Draw conclusions from the research and recommend future research areas to improve DAO governance proposals' summarization further.

Literature:

1. Mahak Gambhir & Vishal Gupta, Recent automatic text summarization techniques: a survey, https://doi.org/10.1007/s10462-016-9475-9
2. Wafaa S. El-Kassas, et. al. Automatic text summarization: A comprehensive survey, https://doi.org/10.1016/j.eswa.2020.113679

Electronically approved by Ing. Magda Friedjungová, Ph.D. on 21 February 2023 in Prague.

FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE

Master's thesis

Machine learning based approach for summarizing governance proposals for decentralized autonomous organizations

Bc. Tiumentsev Herman

Department of Theoretical Computer Science
Supervisor: Alexander Kovalenko

February 15, 2024

Acknowledgements

My heartfelt appreciation extends to my family and friends for their unwavering support, understanding, and encouragement during this academic journey. Their belief in me has been a constant source of motivation.

I am grateful to all the participants who contributed to the research by providing valuable insights and feedback.

Declaration

I hereby declare that the presented thesis is my own work and that I have cited all sources of information in accordance with the Guideline for adhering to ethical principles when elaborating an academic final thesis.

I acknowledge that my thesis is subject to the rights and obligations stipulated by the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular that the Czech Technical University in Prague has the right to conclude a license agreement on the utilization of this thesis as a school work under the provisions of Article 60 (1) of the Act.

In Prague on February 15, 2024 . . . .

Citation of this thesis

Tiumentsev, Herman. *Machine learning based approach for summarizing governance proposals for decentralized autonomous organizations*. Master's thesis. Czech Technical University in Prague, Faculty of Information Technology, 2024.

Abstrakt

Decentralizované autonomní organizace (DAO) se dostávají do popředí jako decentralizované subjekty fungující na základě takzvaných smart contractu a technologie blockchain. Složitost návrhů na řízení v rámci DAO však představuje výzvu pro dostupnost a účast na rozhodovacích procesech. Tato práce řeší problém omezené dostupnosti a účasti tím, že vyvíjí a vyhodnocuje personalizovaný systém založený na strojovém učení pro shrnutí návrhů na správu DAO. Mezi cíle patří prozkoumání současných struktur správy DAO a rozhodovacích procesů, identifikace problémů při sumarizaci návrhů, vyhodnocení různých přístupů k sumarizaci a vývoj přizpůsobeného sumarizačního systému. Cílem systému je zvýšit dostupnost a účast tím, že bude poskytovat stručné a srozumitelné shrnutí návrhů na správu DAO. K posouzení účinnosti systému se používají hodnotící metriky, jako je přesnost, srozumitelnost a relevance. Výsledky ukazují zlepšení přístupnosti, což zdůrazňuje význam specializovaných shrnujících systémů pro zlepšení rozhodovacích procesů v rámci DAO.

Klíčová slova DAO, shrnutí textu, strojové učení, návrhy na správu, NLP

Abstract

Decentralized Autonomous Organizations (DAOs) are gaining prominence as decentralized entities operating on smart contracts and blockchain technology. However, the complexity of governance proposals within DAOs poses challenges to accessibility and participation in decision-making processes. Thesis addresses the problem of limited accessibility and participation by developing and evaluating a personalized machine learning-based system for summarizing DAO governance proposals. The goals include exploring current DAO governance structures and decision-making processes, identifying challenges in summarizing proposals, evaluating different summarization approaches, and developing a customized summarization system. The system aims to enhance accessibility and participation by providing concise and understandable summaries of DAO governance proposals. Evaluation metrics such as accuracy, comprehensibility, and relevance are used to assess the system's effectiveness. Results indicate improvements in accessibility, highlighting the importance of tailored summarization systems in enhancing decision-making processes within DAOs.

Keywords DAO, text summarization, machine learning, governance proposals, NLP

Contents

Introduction **1**

Motivation . . . 1

Problem statement . . . 1

Goals of thesis . . . 2

1 DAOs analysis **3**

1.1 Overview of DAOs . . . 3

1.1.1 Definition and characteristics . . . 3

1.1.2 Growth of DAOs . . . 4

1.1.3 Types of DAOs . . . 5

1.1.3.1 Investment DAOs . . . 5

1.1.3.2 Service DAOs . . . 5

1.1.3.3 Social DAOs . . . 6

1.1.3.4 Protocol DAOs . . . 6

1.1.3.5 Platform DAOs . . . 7

1.1.3.6 Grant DAOs . . . 7

1.1.3.7 In general . . . 8

1.2 Decision-making processes . . . 8

1.2.1 Proposal submission . . . 8

1.2.2 Discussion . . . 8

1.2.3 Voting . . . 8

1.2.3.1 Continuous approval voting . . . 9

1.2.3.2 Optimistic governance . . . 9

1.2.3.3 Delegation . . . 9

1.2.3.4 NFT-based voting . . . 10

1.2.3.5 Quadratic voting . . . 10

1.2.3.6 Token-based voting . . . 10

1.2.3.7 Quorum . . . 10

1.2.3.8 Voting thresholds . . . 10

1.2.3.9 Timed voting 11
1.2.4 Smart contracts 11
1.2.5 In general 11
1.3 DAO problems 11
1.4 Summarizing challenges and limitations 14
1.4.1 Complexity of proposals 14
1.4.2 Complexity of governance process 14
1.4.3 Lack of standardization 15
1.4.4 Technical jargon 15
1.4.5 Limited availability of data 15
1.4.6 Cultural and language Barriers 15
1.4.7 Machine learning limitations 15
1.4.7.1 Contextual understanding 16
1.4.7.2 Tone and intent 16
1.4.7.3 Lack of diversity 16
1.4.7.4 Data privacy 16
1.4.7.5 Human oversight 17
1.4.7.6 In general 17
1.5 Overcoming challenges suggestions 17
1.5.1 Develop standardized formats 17
1.5.2 Utilize domain experts 17
1.5.3 Address cultural and linguistic differences 18
1.5.4 Consider decentralized nature of DAOs 18
1.5.5 Note lack of Standardization 18
1.5.6 Challange machine learning limitations 18
1.5.7 In general 18
1.6 Community survey 18
1.6.1 How often do you vote on governance proposals? 19
1.6.2 How often do you read governance proposals? 19
1.6.3 How do you prefer to read governance proposals? 19
1.6.4 Browsing the governance forum 20
1.6.5 Communication with other members 20
1.6.6 Age of participants 21
1.6.7 How long have you been interested in participating in DAO? 21
1.6.8 How to make community members life easier? 21
1.7 Role of summarization 22
1.7.1 Increased Accessibility 22
1.7.2 Increased Participation 23
1.7.3 Improved Efficiency 23
1.7.4 Increased Consistency 23
1.7.5 Clarity 23
1.7.6 In general 23

2 Summarization methodology **25**
2.1 Data collection . . . 25
2.1.1 Identify sources of data . . . 25
2.1.2 Gather and organize the data . . . 25
2.1.3 Preprocess the data . . . 27
2.1.4 Label the data . . . 27
2.1.5 Multilingual Text Processing . . . 27
2.1.6 Tokenize the data . . . 27
2.2 Exploratory data analysis . . . 27
2.2.1 Topic modeling . . . 29
2.2.1.1 LDA . . . 30
2.2.1.2 Optimal number of topics . . . 32
2.3 Criteria for evaluating the effectiveness of summarization system 34
2.3.1 General criteria . . . 34
2.3.2 ROUGE . . . 34
2.3.2.1 ROUGE-N . . . 35
2.3.2.2 ROUGE-L . . . 35
2.3.3 BERTScore . . . 35
2.3.4 BLEU . . . 36
2.3.5 METEOR . . . 36
2.3.6 GRAD . . . 38
2.3.7 Pyramid . . . 38
2.3.8 NIST . . . 39
2.4 Summarization approaches . . . 40
2.4.1 Types of summarization . . . 40
2.4.2 Advantages and disadvantages . . . 40
2.5 SOTA models overview . . . 41
2.5.1 GPT . . . 41
2.5.1.1 Fine-tuning . . . 42
2.5.1.2 GPT3.5 . . . 42
2.5.1.3 GPT4 . . . 42
2.5.2 BART . . . 42

3 Design **43**
3.1 System design . . . 43
3.2 Frontend . . . 43
3.3 UI/UX design . . . 44
3.4 Application . . . 44

4 Results and discussion **47**
4.1 Dataset collection . . . 47
4.2 EDA results . . . 47
4.2.1 Wordclouds . . . 47
4.2.2 Analyze relevant feature distributions . . . 47

4.2.2.1 Source documents data distributions 47
4.2.2.2 Ground truth summaries data distributions . . 47
4.2.2.3 Categorizing and POS tagging words 47
4.2.2.4 Use of stopwords and punctuations 48
4.2.3 Topic modeling 48
4.2.3.1 Optimal Number of Topics for LDA 48
4.2.3.2 Visualization 48
4.2.3.3 Topic analysis 48
4.3 Generating summaries 49
4.3.1 GPT finetuning 49
4.4 Algorithms evaluation 49
4.4.1 ROUGE, ROUGE aggregated, Bert score, BLEU, METEOR 49
4.4.2 G-eval inspired metric 49
4.4.2.1 Evaluation prompt template 52
4.4.2.2 Metric 1: Relevance 54
4.4.2.3 Metric 2: Coherence 55
4.4.2.4 Metric 3: Consistency 56
4.4.2.5 Metric 4: Fluency 56
4.4.2.6 Results 56

Conclusion and future work **57**

Bibliography **59**

A Acronyms **63**

B Contents of enclosed CD **65**

List of Figures

1.1 Architectural layers for smart contracts (BlockchainHub, 2018) [1] 12
1.2 Voting participation, where * is total number of holders, according to Erherscan 13
1.3 **Voting frequency** on governance proposals 19
1.4 **Reading frequency** of governance proposals 20
1.5 **Preferred methods for reviewing** governance proposals 21
1.6 Subfigures of survey results 22

2.1 Demonstration of the interface from where the data is collected 26
2.2 The process of LDA 29
2.3 GPT models comparision 41

3.1 Backend 43
3.2 Frontend 44
3.3 UI/UX design 45
3.4 Public beta application design 46

4.1 Wordclouds 47
4.2 Source data distributions 48
4.3 Ground truth summaries data distributions 49
4.4 POS tagging words distributions 50
4.5 Stopwords and punctuation distributions 51
4.6 Optimal number of topics metrics for LDA model 51
4.7 Topic Modeling with Gensim (Python) 52
4.8 Finetuning process 52
4.9 ROUGE, ROUGE aggregated, Bert score 54
4.10 BLEU, METEOR 55
4.11 Custom metric inspired by G-Eval 56

List of Tables

2.1 Description of the preprocessing techniques [2] 28

4.1 Named topic labels 53

Introduction

Motivation

Decentralized Autonomous Companies (DAOs) are becoming increasingly popular as a means of organizing and operating companies in a decentralized manner. As a result, there is a growing need for effective text summarization techniques that can help individuals understand complex proposals related to DAO management. This project aims to develop an autonomous text summarization model that can extract the key information from voluminous documents related to DAO management and provide a concise summary to increase the number of participants in voting for certain changes in the course of the development of DAO.

Problem statement

The problem statement for this thesis is the lack of accessibility and participation in the decision-making process of decentralized autonomous organizations (DAOs) due to the complexity of governance proposals. DAOs are decentralized entities that operate based on smart contracts and blockchain technology, and their decision-making process is based on a consensus mechanism where stakeholders participate in decision-making by voting on proposals. However, governance proposals related to DAOs can be complex and lengthy, which can hinder stakeholders' ability to access and participate in the decision-making process.

The problem is further compounded by the limited availability of tools and techniques for summarizing complex DAO governance proposals effectively. Extracting key information and summarizing it concisely and understandably can help increase accessibility and participation in the decision-making process, but existing summarization approaches may not be suitable for DAO governance proposals. Therefore, there is a need for a personalized system that can effectively summarize DAO governance proposals, increasing accessibility and participation in the decision-making process. This thesis aims to address this problem by developing and evaluating a customized machine learning-based system for summarizing DAO governance proposals.

Goals of thesis

The thesis aims to contribute to the field of DAO governance and decision-making processes by developing a personalized system for summarizing proposals that can increase accessibility and participation in the decision-making process. The goals of the thesis are the following:

- Explore the current state of decentralized autonomous organizations (DAOs), their governance structures, and the decision-making processes involved.
- Identify the challenges and limitations related to summarizing proposals related to DAO governance and how they impact accessibility and participation in the decision-making process.
- Evaluate different summarization approaches, such as abstractive and extractive summarization, and determine their advantages and disadvantages in the context of DAO governance proposals.
- Determine the criteria for evaluating the effectiveness of a summarization system for DAO governance proposals, such as accuracy, comprehensibility, and relevance.
- Develop a customized system for summarizing DAO governance proposals that addresses the limitations of existing approaches and meets the identified criteria for effectiveness.
- Test the system on a sample of DAO governance proposals and evaluate its effectiveness using the identified criteria.
- Compare the results of the customized machine learning-based system with existing approaches to summarize DAO governance proposals to demonstrate the improvements in accessibility and participation in the decision-making process.
- Draw conclusions from the research and recommend future research areas to improve DAO governance proposals' summarization further.

Chapter 1

DAOs analysis

1.1 Overview of DAOs

A Decentralized Autonomous Organization (DAO) is an organization that operates autonomously using smart contracts on a blockchain network. DAOs are decentralized, meaning they do not have a central authority or intermediary to manage the organization. Instead, the organization's rules and regulations are encoded in smart contracts that run on a blockchain network, which is a distributed ledger technology [3].

1.1.1 Definition and characteristics

DAOs operate through the consensus of their members, who are known as token holders. Token holders have a say in the decision-making process of the DAO by voting on proposals and allocating resources to various initiatives. Additionally, DAOs use their native cryptocurrency as a means of payment and to incentivize participation.

The following are the key characteristics of DAOs [4]:

Decentralized: DAOs are decentralized, meaning that they do not rely on a central authority to govern their operations. Instead, decision-making is distributed among token holders, who have a say in the organization's affairs.

Autonomous: DAOs operate autonomously, meaning that they execute the rules and regulations encoded in their smart contracts automatically, without requiring human intervention.

Transparent: DAOs operate on a public blockchain, which makes their operations transparent and visible to anyone who wishes to view them.

Immutable: Transactions that occur on a blockchain network are immutable, meaning that they cannot be altered or deleted. Therefore, once a decision is made in a DAO, it is final and cannot be changed.

Democratic: DAOs use a democratic decision-making process, where token holders have a say in the organization's affairs by voting on proposals and allocating resources.

Incentivized: DAOs use their native cryptocurrency as a means of payment and incentivize participation in the decision-making process.

1.1.2 Growth of DAOs

The popularity of DAO tends to grow as the market expands. During this period of growth, new tools, more users, and structural changes within DAOs will become absolutely necessary.

According to the search results, DAOs have seen significant growth in recent years. Here are some key points about the growth of DAOs:

- The total value locked in DAOs increased from around $50 million in early 2020 to more than $5 billion in early 2022 [5].
- Between 2019 and 2020, the number of DAOs grew by 660% [6]. The number of DAOs has increased by 8.8 times, from 700 in May 2021 to 6,000 in June 2022 [7].
- The number of proposals has increased by 8.5 times, and the number of total votes has increased by 8.3 times over the past 12 months [7].
- The number of DAO members increased by a factor of 130x in 2021, reaching 1.6 million [8]. As of April 2023, the total number of DAO governance token holders is 6.9 million, with 2.1 million active voters and proposal makers [6]. The growth in the number of participants is proportional to the increase in the number of decentralized autonomous organizations (DAOs) and their participation rates [9].
- Six DAOs have a treasury of more than $1 billion: Optimism Collective, Arbitrum One, BitDAO, Uniswap, Polygon, and Gnosis [6].
- The World Economic Forum's January 2023 report on the decentralized autonomous organization (DAO) landscape focuses on operational, technical, governance, and legal challenges facing DAO deployment, indicating the growing importance and complexity of these organizations [8].

Overall, the growth of DAOs has been significant, with an increase in the number of DAOs, proposals, and votes, as well as an increase in the total value locked in DAOs. DAOs have the potential to transform many industries, including finance, healthcare, and supply chain management, and their growth is expected to continue in the future.

1.1.3 Types of DAOs

There are different types of Decentralized Autonomous Organizations (DAOs), each with its own purpose and approach to governance and decision-making processes. The following will list the types of DAOs [5].

1.1.3.1 Investment DAOs

Investment DAOs are focused on investing in different projects, such as startups, real estate, or cryptocurrency. Members of the DAO pool their funds to make investments, and they share in the profits or losses. Investment DAOs can benefit from AI-based solutions that can help them identify investment opportunities and manage their portfolios [10]. Examples:

- MetaCartel Ventures (MCV): a community made up of people who want to invest in new projects built on the Ethereum blockchain.

- LAO: a venture DAO that is a community of investors who collaborate on the investment decision-making process and support the collection of projects they invest in. The LAO differs from traditional venture firms by opening themselves up to collaboration from a broader range of people.

- Flamingo DAO: an investment DAO that is focused on investing in NFTs and other digital assets. It is a community-driven DAO that is committed to supporting the growth of the NFT ecosystem.

1.1.3.2 Service DAOs

Service DAOs are focused on providing a specific service or product to the community. For example, a service DAO could be focused on providing decentralized insurance, identity verification, or dispute resolution services. By summarizing proposals, Service DAOs can quickly identify the key features of the proposal and determine whether it aligns with their service offerings. This can help Service DAOs to make informed decisions and provide better services to their customers [11]. Examples:

- Gnosis Safe Multisig: provides a secure and easy-to-use platform for managing digital assets. It is a decentralized platform that allows users to manage their digital assets without the need for a centralized intermediary.

- DAOstack: provides a platform for decentralized collaboration and decision-making. It is a modular platform that allows users to create and customize their own DAOs.

- Aragon: a service DAO that provides a platform for creating and managing decentralized organizations. It is a modular platform that allows users to create and customize their own DAOs.

1.1.3.3 Social DAOs

Social DAOs are primarily concerned with creating social impact and fostering community development. These organizations often allocate resources towards projects that promote open-source software development, environmental sustainability, or educational initiatives. To facilitate efficient decision-making within social DAOs, the process of summarizing proposals can be employed to extract key information regarding the initiative's societal influence, funding needs, and anticipated results. By utilizing this approach, Social DAOs can make well-informed choices while also assessing the effectiveness of their endeavors. [11]. Examples:

- Gitcoin: focused on funding open-source software development. It is a community-driven platform that allows developers to get paid for their work and helps fund open-source projects.

- MolochDAO: focused on funding Ethereum infrastructure projects. It is a community-driven platform that allows members to pool their funds and make collective decisions on which projects to fund

- RadicalxChange: focused on promoting radical social change through the use of blockchain technology. It is a community-driven platform that allows members to collaborate on projects that promote social change.

1.1.3.4 Protocol DAOs

Protocol DAOs are focused on the development and maintenance of a specific blockchain protocol or platform. Members of the DAO are responsible for making decisions related to the protocol's development, upgrades, and maintenance. Summarization can help Protocol DAOs to quickly identify the key features of the proposal, such as the protocol's development, upgrades, and maintenance requirements. This can help Protocol DAOs to manage the development of blockchain protocols [12]. Examples:

- MakerDAO: focused on creating a stablecoin called DAI. It is a decentralized platform that allows users to create and manage their own stablecoins.

- Uniswap: focused on creating a decentralized exchange for trading cryptocurrencies. It is a community-driven platform that allows users to trade cryptocurrencies without the need for a centralized intermediary.

- Yearn Finance: a suite of yield-optimizing decentralized finance protocols that aims to maximize returns on cryptocurrency by arbitraging different lending platforms in search of the best available yield.

1.1.3.5 Platform DAOs

Platform DAOs are focused on developing and maintaining a specific decentralized platform or application. Members of the DAO make decisions related to the platform's development, upgrades, and maintenance. Decentralized governance and artificial intelligence policy with blockchain-based voting can be used to manage the development of decentralized platforms [13]. Examples:

- Compound: focused on creating a decentralized lending platform. It is a community-driven platform that allows users to lend and borrow cryptocurrencies without the need for a centralized intermediary.

- Aave: focused on creating a decentralized lending and borrowing platform. It is a community-driven platform that allows users to lend and borrow cryptocurrencies without the need for a centralized intermediary.

- Balancer: focused on creating a decentralized exchange for trading cryptocurrencies. It is a community-driven platform that allows users to trade cryptocurrencies without the need for a centralized intermediary.

1.1.3.6 Grant DAOs

Grant DAOs are focused on funding various initiatives or projects that align with the DAO's mission or vision. Members of the DAO vote on proposals submitted by individuals or organizations seeking funding. Grant DAOs can use proposal summarization to improve voter turnout and engagement. Proposal summarization can help Grant DAOs quickly identify the key features of the proposal, such as the initiative or project seeking funding, funding requirements, and expected outcomes [14]. Examples:

- DAOhaus: focused on funding projects that are building on the Ethereum ecosystem. It is a community-driven platform that allows members to vote on which projects to fund.

- BitDAO: the most substantial decentralized autonomous organization (DAO)-managed fund aimed at fostering the expansion of open finance and supporting the development of decentralized token-driven economies.

- Global Coin Research: community-first research and investment DAO that is focused on Web3. The DAO also invests in transformative projects such as Lens Protocol, which aims to create a decentralized, user-centric social media space.

1.1.3.7 In general

Overall, the different types of DAOs are characterized by their focus and approach to governance and decision-making. By understanding the different types of DAOs, it is possible to appreciate their versatility and potential impact in different domains.

1.2 Decision-making processes

Decision-making processes in DAOs are designed to ensure that decisions are made in a transparent, democratic, and efficient manner. The following will be considered some of the key aspects of decision-making processes in DAOs [1].

1.2.1 Proposal submission

Any member of a DAO can submit a proposal for consideration. Proposals can range from allocating funds to a new project to making changes to the organization's rules and regulations. Proposal submission is an important aspect of DAO governance as it allows members to voice their opinions and suggest changes to the organization.

1.2.2 Discussion

Once a proposal is submitted, members of the DAO can discuss the proposal and provide feedback. This helps to ensure that all perspectives are heard and that the proposal is thoroughly evaluated before a decision is made. The discussion can take place on various platforms, such as forums, chat rooms, or social media. The discussion should be respectful, constructive, and focused on the proposal's merits.

1.2.3 Voting

After a proposal has been discussed, it is put to a vote. Members of the DAO can vote either in favor of or against the proposal, and the proposal is accepted or rejected based on the outcome of the vote. The voting process should be transparent, secure, and accessible to all members. The DAO should use a reliable and decentralized voting system to ensure that the vote is fair and accurate. DAO voting processes balance efficiency and effectiveness considerations. They seek to avoid familiar problems in governance systems, such as rational apathy (where voters do not participate because it requires time and effort, but each voter has a minimal impact on outcomes) and plutocracy (a concentration of power deriving from wealth). Token-based quorum voting is the simplest form of voting, used in many leading DAOs, including Uniswap and Compound. For a proposal to be submitted and passed, a certain number or percentage of tokens must participate. Selecting the proper quorum requirements can be challenging. DAOs have developed alternative approaches to quorum thresholds. Some DAOs support the delegation of voting or proposal power to others through a representative system, while others provide greater voting power to individuals who "lock up" or stake their tokens in an escrow smart contract for a fixed amount of time. While DAO frameworks support a wide variety of voting processes, many DAOs opt to implement complex voting practices off-chain through tools including Snapshot, Discourse, and Commonwealth [12][15].

1.2.3.1 Continuous approval voting

Continuous approval voting allows new proposals to be submitted at any time, as long as they surpass the voting weight of the last successful proposal implemented. The more votes there are on the system's current state, the more secure the system is from any "rogue" proposals. However, this can make it harder for proposals to overcome the status quo [12].

1.2.3.2 Optimistic governance

Optimistic governance attempts to reduce voter fatigue by dramatically lowering the number of proposals upon which a token holder is expected to vote. This model assumes that proposals pass unless there is a strong objection, requiring a rejection threshold rather than an approval quorum. The rejection threshold is usually much lower than a typical approval quorum, meaning more voices may be heard in this process. However, this model still relies on active monitoring of proposals and adequate contestation periods to prevent problematic proposals slipping through [12].

1.2.3.3 Delegation

Delegation allows token holders to outsource decision-making and/or direct their proposing and/or voting rights to value-aligned people or groups they trust. Delegation is similar to proxy votes in traditional finance. Delegates can be individuals or other DAOs. Many DAOs elect councils or committees as trusted token-holder representatives. These representatives act as a quasi-board of directors, often elected via a decentralized election process that lacks authority to act without input and support from the broader community. They may be elected on a regular basis or by fulfilling a delegation threshold that runs on a rolling basis. Some representatives step in to vote on proposals only when token holders fail to reach a quorum. Since each council or committee member usually has only one vote, the tendency towards plutocratic decision-making is also reduced. Representatives abusing their position may be removed by token-holder vote [12].

1.2.3.4 NFT-based voting

Several protocols are moving towards NFT (non-fungible token [16]) based voting that moves away from a one-token, one-vote paradigm towards a one-person, one-vote model. NFT-based voting may mitigate the risk of plutocracy common to token-weighted voting that has been well outlined by Vitalik Buterin. DAOs experimenting with NFT-based voting include Optimism, Element Finance, and Marinade Finance [12].

1.2.3.5 Quadratic voting

Quadratic voting is another method that attempts to reduce the tendency towards plutocracy and is employed by Gitcoin DAO. Votes are counted according to their square root, so 100 different token holders voting one token for a proposal will have greater weight than one large holder casting 200 tokens. Quadratic voting systems must address the challenge of Sybil attacks, whereby one actor simply splits their tokens between multiple wallets [12].

1.2.3.6 Token-based voting

In most DAOs, voting is based on the number of tokens that a member holds. This means that members with more tokens have more voting power, which can be a source of controversy in some DAOs. Token-based voting is designed to ensure that members who have invested more in the DAO have more say in the decision-making process. However, it can also lead to centralization and inequality if a small group of members hold a significant number of tokens.

1.2.3.7 Quorum

DAOs require a quorum, which is a minimum number of members who must participate in a vote for it to be valid. This ensures that decisions are made with the participation of a sufficient number of members. The quorum should be set at a reasonable level to ensure that decisions are not delayed or blocked by a small number of members.

1.2.3.8 Voting thresholds

DAOs typically have a voting threshold that must be met before a proposal is accepted or rejected. This threshold can be a certain percentage of the total token supply or a specific number of votes. The voting threshold should be set at a level that ensures that decisions are made with the support of a significant number of members. However, it should not be set too high, as this can lead to a deadlock in the decision-making process.

1.2.3.9 Timed voting

DAOs use timed voting to prevent proposals from being held up indefinitely. Each proposal is given a set amount of time for members to vote, and if the proposal does not meet the required voting threshold within the specified time frame, it is rejected. Timed voting ensures that decisions are made in a timely and efficient manner, and that proposals are not delayed or blocked by a small number of members.

1.2.4 Smart contracts

Smart contracts are used to execute proposals that the DAO has approved. Once a proposal has been approved and the required conditions have been met, the smart contract automatically executes the proposal.

The appeal of smart contracts lies in their ability to operate on a par with external accounts, enabling them to autonomously generate other contracts and send messages [1]. When multiple smart contracts are coordinated within a network, they create a sophisticated autonomous system of functions. Additionally, the Ethereum blockchain serves as the foundational technology for recording transactions and establishing consensus-driven logic. These layers are depicted in Figure 1.1.

1.2.5 In general

Given the challenges outlined above, an increasing number of protocols are trending towards governance minimization or limiting the number of decisions to be made by humans, often via automation at the technical layer. This results in a streamlined governance system, but the more automated governance becomes, the less adaptive the protocol will be, and it is unclear how these protocols will stand the test of time as parameters ossify.

Overall, the decision-making processes in DAOs are designed to promote democratic decision-making, ensure member participation, and maintain the autonomy of the organization. By using token-based voting, quorums, and smart contracts, DAOs are able to make decisions in a transparent and efficient manner, allowing them to operate autonomously and effectively.

1.3 DAO problems

One of the biggest problems in DAOs is information overload, inconsistent information and highly specialized information that waterfalls on users. The

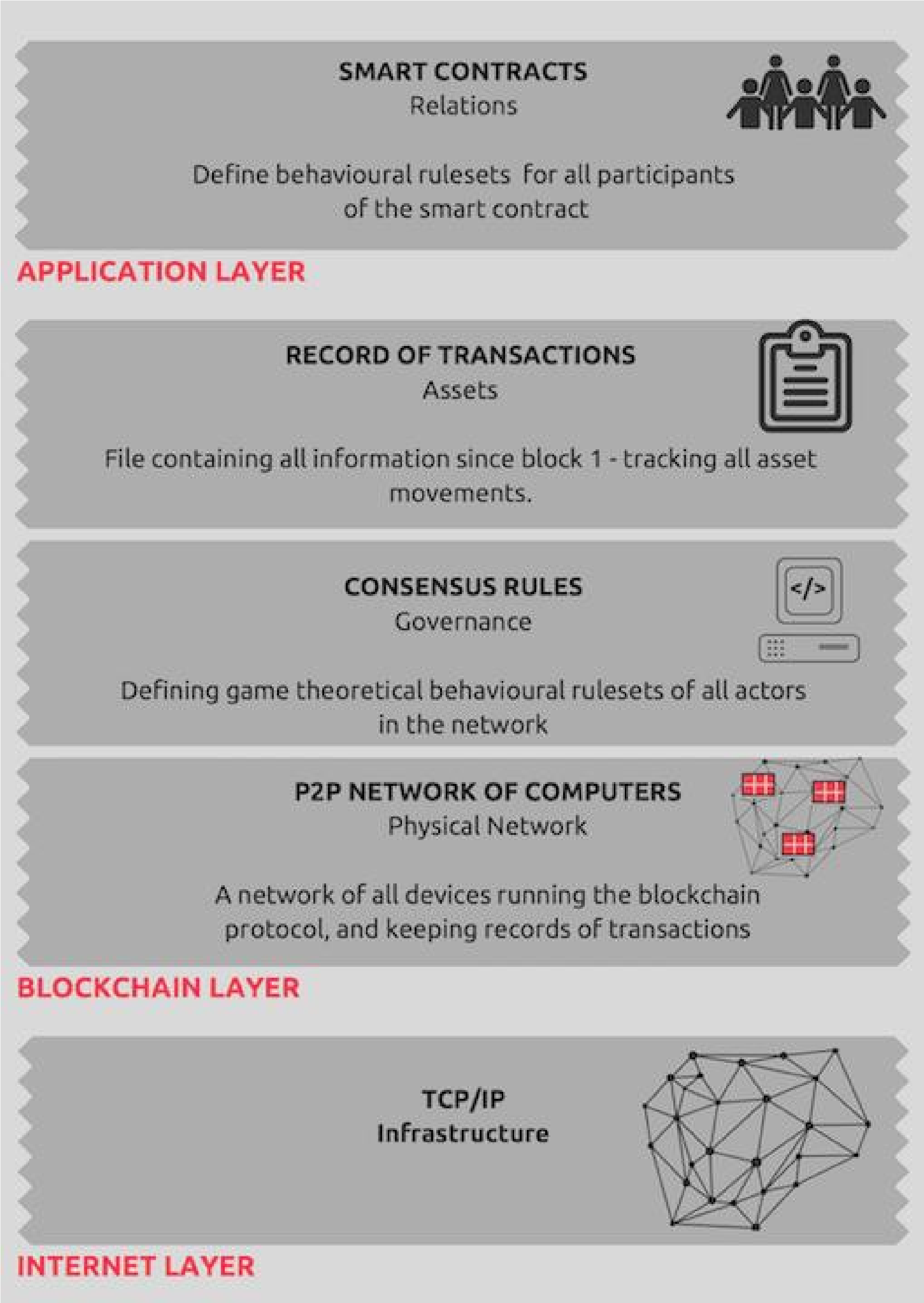


Figure 1.1: Architectural layers for smart contracts (BlockchainHub, 2018) [1]

following is a list of problems that DAOs have:

- Information regarding updates gets lost in the shuffle.

- Participation is extremely hard and time-consuming for newcomers. Voting participation remains low. 2% average of token holders vote (see Figure 1.2).
- Too many sources of information lead to confusion.
- There are thousands of DAOs.
- Highly specialized information requires copious amounts of time to read and understand.
- Every DAO has multiple social media sources.
- Motivation to participate is low and the benefits are unclear.
- People are uncertain of their influence within the DAO.

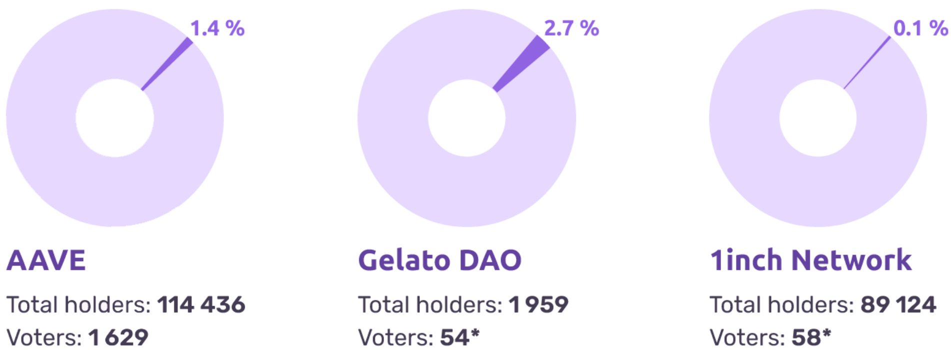


Figure 1.2: Voting participation, where * is total number of holders, according to Erherscan

Despite the potential of DAOs to revolutionize decentralized decision-making, the challenges outlined in the previous section illuminate the existing hurdles that hinder widespread participation and comprehension. The information overload, lack of clarity in benefits, and low motivation to participate are persistent issues that contribute to the meager average of 2% token holder voting participation. Additionally, the sheer number of DAOs, each with its own specialized information spread across multiple sources, exacerbates the confusion and time constraints faced by potential participants. These challenges underscore the need for effective solutions to enhance the accessibility and engagement of individuals within DAOs, such as proposal text summarizing.

1.4 Summarizing challenges and limitations

Summarizing DAO governance proposals can be a challenging task due to the complexity of the proposals and the decentralized nature of DAOs. Following are some of the challenges and limitations related to summarizing DAO governance proposals.

1.4.1 Complexity of proposals

DAO governance proposals often present a formidable challenge due to their inherent complexity and technical nature [17]. These proposals frequently employ specialized language and concepts that may appear unfamiliar to certain members of the DAO community. Effectively summarizing such proposals demands a deep understanding of the subject matter and a judicious selection of the most crucial points and recommendations. The intricacy of these proposals can complicate the process of capturing their nuanced details in a summary, potentially resulting in misunderstandings or oversimplifications.

1.4.2 Complexity of governance process

Simultaneously, the governance process within DAOs introduces its own layer of complexity. It commonly involves the submission of proposals followed by collective voting, culminating in a substantial volume of information to digest [18]. The ability to summarize these proposals accurately and succinctly is paramount for fostering accessibility and promoting active participation. It enables participants to swiftly grasp the salient points and make well-informed decisions.

Moreover, the transparency and autonomy that characterize DAO governance bring forth additional challenges. While transparency is a desirable feature of DAOs, it can inadvertently lead to information overload and the difficult task of extracting the right information from all that flow. [19]. Additionally, the self-governing nature of DAOs means that decision-making processes may transpire without the involvement of centralized entities, thereby posing difficulties in ensuring accountability and inclusivity within these processes [19].

These challenges collectively impact the accessibility and participation levels within the DAO decision-making framework. The intricate and voluminous nature of the information can hinder participants from engaging effectively and making informed choices [18]. Additionally, the transparency and autonomy inherent to DAO governance may inadvertently create barriers for individuals who lack familiarity with the underlying technology or the technical expertise needed to navigate the system effectively [19].

1.4.3 Lack of standardization

The lack of standardization in DAO governance proposals can make summarization difficult because there is no consistent structure or format to follow. This means that summarization tools must be flexible enough to handle a wide variety of proposals, which can be a significant technical challenge [17]. Additionally, without a standardized format for proposals, it can be difficult to compare and evaluate the effectiveness of different summarization methods.

1.4.4 Technical jargon

DAO governance proposals often contain technical jargon and specialized language that may be difficult for non-experts to understand. Model may need to have a deep understanding of the technical concepts and language used in the proposal to be able to accurately summarize its content [17]. Furthermore, the use of technical jargon can also make it challenging to convey the essential points and recommendations of the proposal in a clear and concise manner.

1.4.5 Limited availability of data

There is a limited amount of data available on DAO governance proposals, which can make it difficult to train machine learning models for summarization [20]. Moreover, DAO governance is a relatively new field, and there is still much to be learned about how it works. This can make it challenging to develop effective summarization approaches [11].

1.4.6 Cultural and language Barriers

The cultural and linguistic diversity of DAOs can pose a challenge to summarizing governance proposals. Some members may have limited proficiency in English, which is the primary language used for many proposals. Additionally, proposals may contain cultural references that are unfamiliar to some members, which can make summarization more challenging. Overcoming these barriers requires developing summarization tools that can accurately capture the intended meaning of the proposals, while also being accessible to all members.

1.4.7 Machine learning limitations

Machine learning-based summarization methods may have limitations in summarizing DAO governance proposals. Moreover, the algorithms may be biased towards certain types of proposals or languages, leading to a lack of diversity in the summaries. This section delves into the constraints of machine learning-based summarization approaches when summarizing DAO governance proposals.

1.4.7.1 Contextual understanding

Machine learning-based summarization methods may struggle to understand the context of the DAO governance proposals. The proposals may contain technical terms, jargon, and references to other proposals or documents. Without a deep understanding of the context, the summarization algorithm may produce inaccurate or incomplete summaries. To overcome this limitation, the summarization algorithm should be trained on a large dataset of DAO governance proposals to develop a contextual understanding of the subject matter.

1.4.7.2 Tone and intent

Machine learning-based summarization methods may struggle to understand the tone and intent of the DAO governance proposals. The proposals may contain emotional language, sarcasm, or irony, which may be difficult for the algorithm to detect. Moreover, the proposals may contain hidden agendas or biases, which may be difficult for the algorithm to identify. To overcome this limitation, the summarization algorithm should be trained on a diverse dataset of DAO governance proposals to develop a nuanced understanding of the tone and intent.

1.4.7.3 Lack of diversity

DAO governance proposals often lack standardization, presenting a significant challenge for summarization systems. These proposals may adopt varying formats, structures, and criteria, making it difficult to establish a uniform summarization approach. Furthermore, they may be subject to distinct regulatory frameworks depending on the jurisdiction, further complicating the development of a universally applicable summarization method.

Additionally, machine learning-based summarization methods can introduce bias, both towards certain proposal types and specific languages, potentially resulting in a lack of diversity within the summaries. Biases can emerge if the algorithm is trained on a dataset that does not adequately represent the entire DAO ecosystem, skewing the summaries. Furthermore, such algorithms might struggle when summarizing proposals written in languages other than English.

1.4.7.4 Data privacy

Machine learning-based summarization methods may struggle with data privacy concerns. DAO governance proposals may contain sensitive information, such as financial data or personal information, which may be difficult to summarize without compromising data privacy. Moreover, the summarization algorithm may be vulnerable to cyber-attacks or data breaches, leading to the exposure of sensitive information. To overcome this limitation, the summarization algorithm should be designed with data privacy in mind, and the data should be stored securely.

1.4.7.5 Human oversight

Machine learning-based summarization methods should not be relied upon solely to summarize DAO governance proposals. Human oversight should ensure that the summaries are accurate, complete, and unbiased. The summarization algorithm may be used as a tool to assist humans in summarizing the proposals, rather than replacing them. Moreover, the human summarizer should have a deep understanding of the subject matter and the ability to identify biases and inaccuracies in the summary.

1.4.7.6 In general

Machine learning-based summarization methods have limitations when it comes to summarizing DAO governance proposals. Overcoming these limitations requires a deep understanding of the subject matter, a diverse dataset of DAO governance proposals, and a nuanced understanding of the tone and intent of the proposals. Moreover, data privacy concerns should be addressed, and human oversight should be used to ensure the accuracy and completeness of the summaries.

1.5 Overcoming challenges suggestions

To navigate the challenges outlined earlier, several strategies can be employed to enhance the effectiveness of summarizing DAO governance proposals.

1.5.1 Develop standardized formats

DAOs could work to develop standardized formats for governance proposals that make them easier to summarize. This could include using consistent language and formatting, as well as providing clear headings and subheadings that identify the main points of the proposal.

1.5.2 Utilize domain experts

Summarizers with deep domain expertise in DAO governance could be utilized to help ensure that summaries accurately capture the essential points and recommendations of proposals. These experts could also help develop guidelines for summarizing proposals that take into account the complexity and technical nature of the content.

1.5.3 Address cultural and linguistic differences

To ensure that summaries are accessible to all members of the DAO, it may be necessary to provide translations of summaries into multiple languages. Additionally, summarizers could work to take into account cultural differences when summarizing proposals, such as by avoiding idiomatic expressions that may not be understood by all members.

1.5.4 Consider decentralized nature of DAOs

To overcome this challenge, web scraping tools should be used to gather all the relevant data from various platforms. Natural language processing (NLP) techniques should also help yo translate and summarize proposals written in different languages.

1.5.5 Note lack of Standardization

To overcome this limitation, the summarization algorithm should be trained on a diverse dataset of DAO governance proposals that represent the entire DAO ecosystem.

1.5.6 Challange machine learning limitations

To overcome this challenge, the summarization system should use a hybrid approach that combines machine learning-based algorithms with human oversight. To improve accuracy and completeness, the machine learning algorithm should be trained on a diverse dataset of DAO management proposals. It should also use a human summarizer to review and edit the summary produced by the machine learning-based algorithm.

1.5.7 In general

Working with the challenges related to summarizing DAO governance proposals requires a systematic approach that addresses each challenge. By developing a deep understanding of the subject matter, using web scraping tools and machine learning-based algorithms, developing a standardized format, using a diverse team of summarizers, and using a hybrid approach that combines machine learning-based algorithms with human oversight, we can overcome these challenges and produce accurate, complete, and unbiased summaries of DAO governance proposals.

1.6 Community survey

In collaboration with our beta testing initiative, my colleagues from the Holdim team and I sought participation from community members to complete a concise questionnaire. We have gathered approximately 200 responses.

1.6.1 How often do you vote on governance proposals?

The gathered data indicates (see Figure 1.3) that actual participation in the voting process is notably influenced by the presence of active voting opportunities and the appeal of the proposals themselves. This insight underscores the need for an efficient summarization system that can help surface relevant and interesting proposals to community members, potentially increasing their engagement in the voting process.

Actual participation in the voting process is heavily influenced by the likelihood of seeing active voting and proposals that pique the interest of the community.

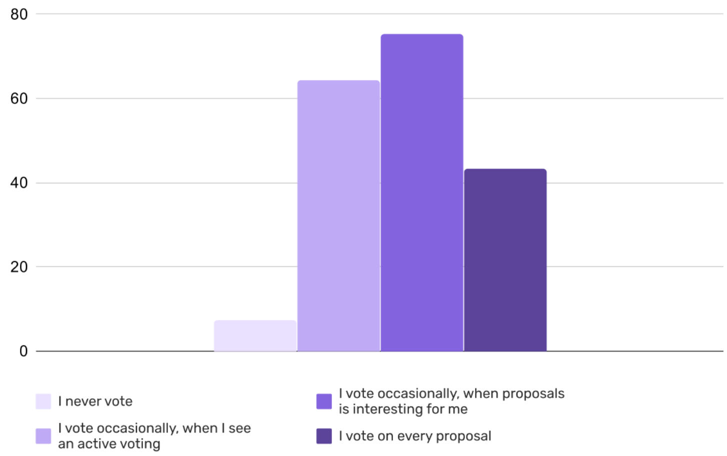


Figure 1.3: **Voting frequency** on governance proposals

1.6.2 How often do you read governance proposals?

The frequency with which community members read governance proposals is closely related to their participation in DAO decision-making. The data reveals (see Figure 1.4) that participants often rely on the availability of time and the appeal of specific proposals as deciding factors for reading them.

So dependence on the chances can be seen in the frequency of reading the actual proposals. Most participants read governance proposals occasionally when they have the time and may find some that are of interest to them.

1.6.3 How do you prefer to read governance proposals?

Most respondents typically only focus their time on proposals that seem interesting to them. They usually determine the worth of their attention by

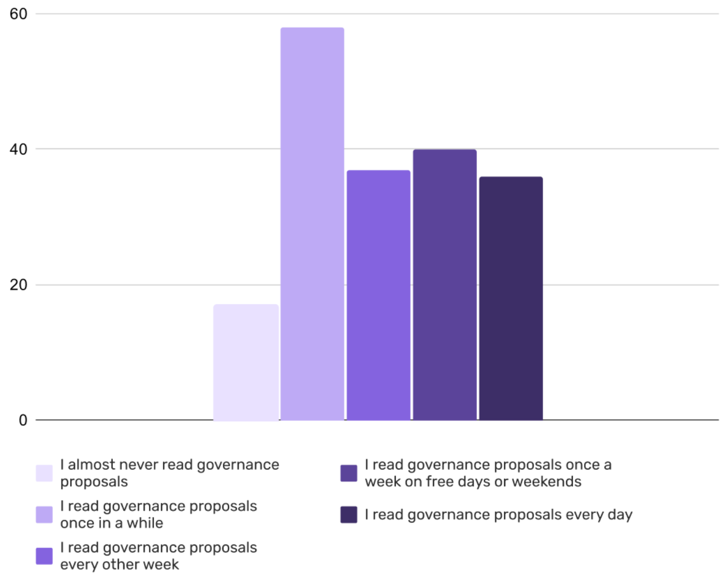


Figure 1.4: **Reading frequency** of governance proposals

simply looking at the title of the proposal. Additionally, a significant portion of people quickly scan the text of the proposal to gain an understanding of its contents (see Figure 1.5). This preference for rapid evaluation implies the necessity of a summarization system that can capture the essence of proposals succinctly. Such a system would aid community members in their initial assessment of proposal relevance and potentially encourage them to delve deeper into the details.

1.6.4 Browsing the governance forum

Participants tend to spend more time browsing the governance forums (see Figure 1.6a), although this may not be the most effective use of their time as they need to open each forum page individually and seek out topics of interest. This finding underscores the need for a summarization that can aggregate and present forum discussions and proposals in a manner that saves community members' time and directs their attention to the most pertinent content.

1.6.5 Communication with other members

The data suggests that community members frequently engage with each other (see Figure 1.6b), emphasizing the importance of community interaction within DAOs. A machine learning-based summarization system could potentially facilitate these interactions by improving the accessibility and understanding of governance proposals, enabling more informed discussions and decision-making.

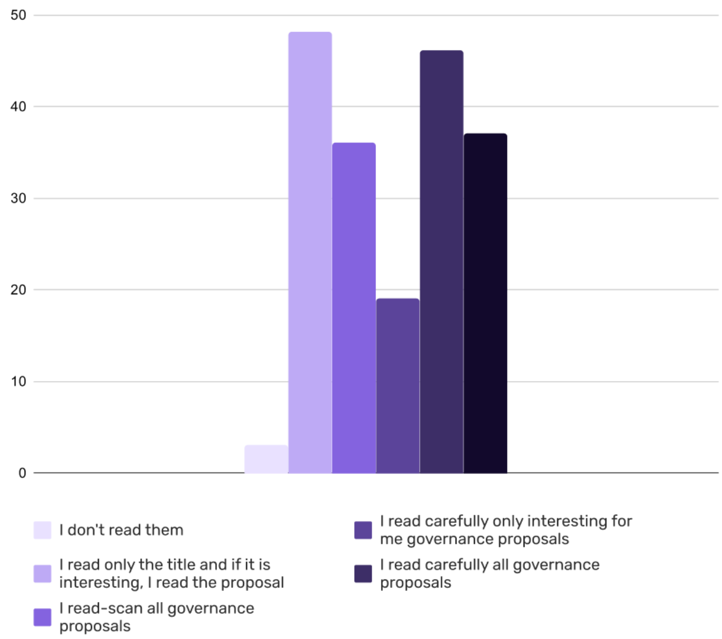


Figure 1.5: **Preferred methods for reviewing** governance proposals

1.6.6 Age of participants

The demographic information gathered reveals that the majority of participants are relatively young, with a significant portion being aged between 18 and 35 years old (see Figure 1.6c). Additionally, a considerable number of respondents are new to the concept of DAOs. This demographic insight informs the design of the summarization system, suggesting the need for user-friendly interfaces and explanations that cater to a potentially less-experienced audience.

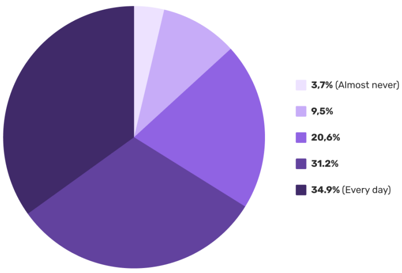
1.6.7 How long have you been interested in participating in DAO?

Half of the respondents are relatively new to the space of DAOs (see Figure 1.6d), with most having just recently become interested or have had an interest for about a year.

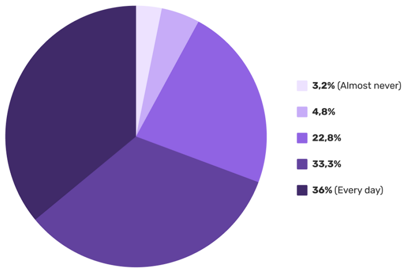
1.6.8 How to make community members life easier?

We asked for ideas on how make DAO life of community members easier. The following are ideas from people in the community.

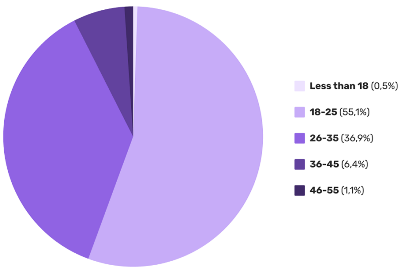
1. *Improve exploration:* Many participants in the survey found it difficult to comprehend the current explanations, visions, and aspirations of DAOs.



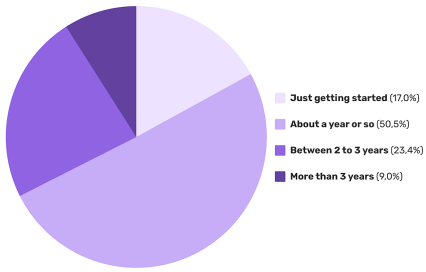
(a) **Viewing governance forums frequency**



(b) **Frequency of communication among community members**



(c) **Age distribution of participants**



(d) **Duration of interest in participating in DAOs**

Figure 1.6: Subfigures of survey results

2. *Access to community:* Additionally, it can be hard for people to identify ways to get involved, so it is important to make the interfaces easier to understand and use so that people can delegate and vote on proposals with ease.

3. *Security and incentives for community:* Lastly, respondents suggested that the security of the DAOs should be strengthened, as well as incentives for the community.

1.7 Role of summarization

Summarizing DAO governance proposals can have several benefits for the accessibility and participation of members in the decision-making process. Some of them are presented below.

1.7.1 Increased Accessibility

Summarizing proposals can make them more accessible to members who may not have the time or expertise to read and understand lengthy and complex documents. Summaries can provide a quick and accessible overview of the key points and recommendations, which can help members make informed decisions.

1.7.2 Increased Participation

Summarizing proposals can also increase members' participation in the decision-making process by making the proposals more easily digestible. Members may be more likely to engage with the proposals if they are presented in a summarized format that is easy to understand and follow.

1.7.3 Improved Efficiency

Summarizing proposals can improve the efficiency of the decision-making process by reducing the time and effort required to review and analyze proposals. Summaries can provide a quick overview of the key points and recommendations, allowing members to focus their attention on the most important aspects of the proposals.

1.7.4 Increased Consistency

Summarizing proposals can also improve the consistency of decision-making by providing a standardized format for evaluating proposals. Summaries can help ensure that all members are considering the same key points and recommendations, which can lead to more consistent decision-making outcomes.

1.7.5 Clarity

Proposals in DAO governance can be lengthy and complex, with technical jargon that can be difficult to understand for those who are not experts in the field. Summaries can provide a clear and simple overview of the proposal, allowing members to quickly understand the proposal's purpose and goals.

1.7.6 In general

Overall, summarizing DAO governance proposals can help ensure that all members have access to the information they need to make informed decisions, which can lead to more equitable and effective decision-making outcomes.

Chapter 2

Summarization methodology

2.1 Data collection

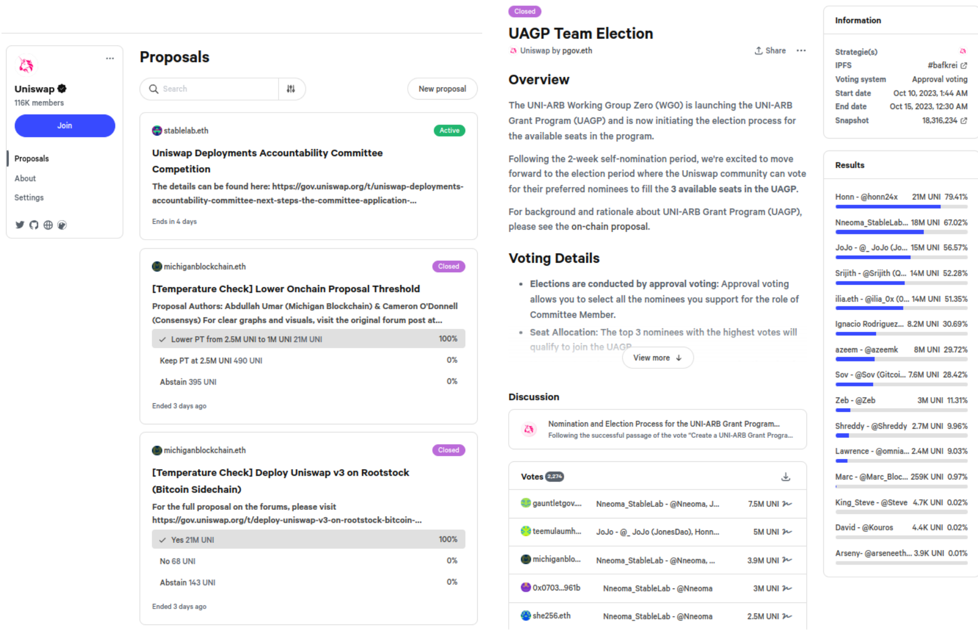
Data collection and preprocessing process will be crucial to the success of the research. Following are some steps that can be taken in this process [2]:

2.1.1 Identify sources of data

The first step is to identify the publicly available sources of data that will be used to develop and test the system. **Snapshot** [21] is a decentralized governance platform for the Ethereum ecosystem that allows token holders to participate in the decision-making process DAOs. It is a simple and gas-efficient way to vote on proposals and delegate voting power to trusted parties. Snapshot is designed to be flexible and can be used by any DAO that wants to implement a governance system. It is also open-source, meaning that anyone can contribute to its development and use it for their own projects. Figure 2.1 demonstrates what snapshots look like in this portal.

2.1.2 Gather and organize the data

Once the sources of data have been identified, the next step is to gather and organize the data. To collect data from snapshot.org, a **GraphQL API** [22] was utilized. This API allows for efficient and flexible data retrieval by enabling clients to specify the exact data they need in a single request. By using GraphQL, the necessary information could be obtained without the need for multiple API calls or excessive data transfer. The snapshot.org GraphQL API likely provides a well-documented schema, allowing one to understand the available data and its structure.



(a) Snapshots of one company listed in sight of the list

(b) Detailed inspection of a single snapshot

Figure 2.1: Demonstration of the interface from where the data is collected

```
query {
  proposals (
    first: 20,
    skip: 0,
    where: {
      space_in: ["yam.eth"],
      state: "closed"
    },
    orderBy: "created",
    orderDirection: desc
  ) {
    id
    title
    body
    choices
    start
    end
    snapshot
    state
    scores
    scores_by_strategy
    scores_total
    scores_updated
    author
    space {
      id
      name
    }
  }
}
```

Listing 2.1: GraphQL Query Example

2.1.3 Preprocess the data

Preprocessing is an essential step to ensure that the data is clean and ready for use and analysis. This may involve removing unnecessary information, such as metadata, formatting the text to ensure consistency, and correcting any spelling or grammatical errors. The techniques presented are detailed in the Table 2.1.

2.1.4 Label the data

To train a supervised machine learning model, the data needs to be labeled with summary information. For this study, the dataset was created as follows: it was based on the GPT3.5-turbo model. Several samples were generated and the best one was selected. Then these generalizations were corrected, supplemented and added by humans.

2.1.5 Multilingual Text Processing

As DAOs can be global with members from diverse linguistic backgrounds, multilingual text processing techniques can be used to handle proposals written in different languages. This can involve techniques such as machine translation, language detection, and named entity recognition for multiple languages.

2.1.6 Tokenize the data

Tokenization is the process of breaking the text into smaller units, such as words or phrases. This step is essential for the machine learning model to understand the text and generate summaries accurately.

2.2 Exploratory data analysis

Exploratory data analysis (EDA) is a crucial step in developing system for summarizing proposals [25] related to DAO governance. EDA involves analyzing the dataset to identify its characteristics and anomalies, which can inform the development of an effective summarization system. By examining the dataset, EDA can help to identify the most relevant features for summarization, determine the appropriate summarization approach to use, and evaluate the effectiveness of the developed summarization system. EDA is essential for ensuring the accuracy, relevance, and comprehensibility of the generated summaries, and for improving accessibility and participation in the decision-making process. EDA is crucial for several reasons.

Firstly, it allows to understand the nature of the proposals related to DAO governance. By analyzing the dataset, key features such as structure, content, and language used in these sentences can be identified. This understanding is

№	Technique name	Description
1	Lower text	Lowercasing the text primarily serves to ensure that words like ”Hello” and ”hello” are not treated as distinct entities, as they represent the same word. This practice reduces the volume of words that need to be stored in the dictionary simultaneously [23].
2	Removing @mention	Eliminating @mentions aids in excluding user references that do not contribute pertinent information to the text's sentiment analysis.
3	Removing URL	The removal of URLs encompasses the elimination of any web addresses from the tweet, encompassing URLs starting with HTTP, https, and even pic:\\ (denoting image URLs).
4	Removing punctuation	Punctuation and non-alphanumeric characters are extracted from the original text.
5	Removing the hashtag	Hashtags are extracted from the text and stored in a separate column, with their significance potentially utilized in subsequent processes.
6	Removing whitespace	Whitespace is omitted from the text as it lacks semantic meaning, simplifying computational operations.
7	Tokenization	Tokenization entails breaking down each sentence into individual words or text units.
8	Removing encoded text formats	In this study, encoded text formats are purged, retaining only those that convey specific meanings. This includes the removal of terms such as xbf and x9a, among others.
9	Removing stop words	Stop words, such as ”a,” ”an,” and ”the,” are eliminated since they contribute little semantic value to the text. This enhances the accuracy of sentiment analysis, utilizing the remaining meaningful text [24].
10	Stemming	Stemming involves reducing words to their root meanings, reducing the total word count and enhancing computational efficiency.

Table 2.1: Description of the preprocessing techniques [2].

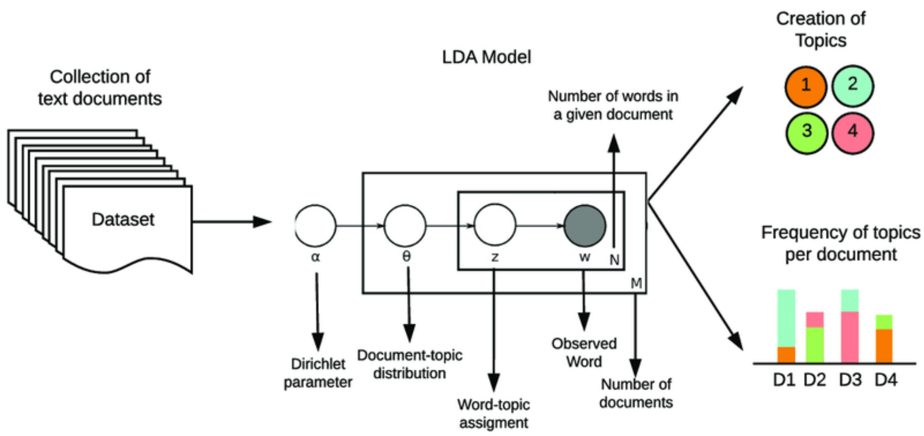


Figure 2.2: The process of LDA

essential for developing an effective summarization system that can accurately capture the important information and convey it in a concise manner [26].

Secondly, EDA helps in identifying the current challenges and limitations related to summarizing proposals related to DAO governance. By analyzing the dataset, common problems such as the presence of technical jargon, complex sentence structures, and redundant information can be identified. These challenges can impact the accessibility and participation in the decision-making process, as they may deter individuals from engaging with the proposals. Understanding these challenges is crucial for developing a customized system that addresses these limitations and improves accessibility [27].

Thirdly, EDA allows for the evaluation of different summarization approaches, such as abstractive and extractive summarization, in the context of DAO governance proposals. By analyzing the dataset and applying different summarization methods, it is possible to compare the advantages and disadvantages of each approach. This evaluation is important for determining the most effective method for summarizing complex documents [28].

Exploratory data analysis helps in understanding the nature of the proposals, identifying challenges and limitations, evaluating different summarization approaches. By conducting a thorough EDA, the groundwork can be laid for developing an effective and accessible debriefing system that promotes greater participation in DAO management decision-making.

2.2.1 Topic modeling

Topic modeling can be used to identify the main themes and topics present in a set of DAO governance proposals. This can help to guide the summarization process by identifying the most important topics and ensuring that the summary captures the main points related to these topics.

2.2.1.1 LDA

Latent Dirichlet Allocation (LDA) is a widely used method for topic modeling in natural language processing. It is a generative probabilistic model that assumes that each document in a corpus is a mixture of a small number of topics, and each topic is a probability distribution over words [29].

The LDA model assumes that each document is generated as follows:

- Choose a distribution over topics from a Dirichlet distribution with parameter α.
- For each word in the document:
  1. Choose a topic from the distribution over topics.
  2. Choose a word from the topic's distribution over words from a Dirichlet distribution with parameter β.

The goal of LDA is to infer the topic distribution for each document and the word distribution for each topic. This is done using Bayesian inference, specifically variational inference or Gibbs sampling.

The output of LDA is a set of topics, each represented as a probability distribution over words. These topics can be interpreted by examining the most probable words in each topic. The topic distribution for each document can also be examined to understand the main themes present in the document.

The mathematical formulas related to LDA for topic modeling are as follows:

- α: The parameter of the Dirichlet prior on the per-document topic distributions.
- β: The parameter of the Dirichlet prior on the per-topic word distributions.
- K: The number of topics.
- N: The number of words in a document.
- M: The number of documents in the corpus.
- $w_{i,j}$: The jth word in the ith document.
- $z_{i,j}$: The topic assigned to the jth word in the ith document.

The joint distribution of the LDA model is given by:

$$p(\boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\theta}, \boldsymbol{\phi}|\alpha, \beta) = \prod_{i=1}^{M} p(\boldsymbol{\theta}_i|\alpha) \prod_{j=1}^{N} p(z_{i,j}|\boldsymbol{\theta}_i) p(w_{i,j}|z_{i,j}, \boldsymbol{\phi}_{1:K}) p(\boldsymbol{\phi}_k|\beta) \quad (2.1)$$

where $\boldsymbol{\theta}_i$ is the topic distribution for the ith document, $\boldsymbol{\phi}_k$ is the word distribution for the kth topic, and $\boldsymbol{w}$ and $\boldsymbol{z}$ are the observed words and topic assignments, respectively.

The E-step (Expectation step) of the EM (Expectation-Maximization) algorithm for LDA involves computing the posterior distribution of the topic assignments given the observed words and the current estimates of the topic and document distributions. The M-step (Maximization step) involves updating the estimates of the topic and document distributions based on the posterior distributions computed in the E-step (Expectation step).

2.2.1.1.1 E-step The E-step [29] (Expectation step) of the EM (Expectation-Maximization) algorithm is the first step in each iteration of the algorithm. In this step, the algorithm computes the expected value of the complete-data log-likelihood function, given the observed data and the current estimates of the model parameters.

The complete-data log-likelihood function is the log-likelihood function of the joint distribution of the observed data and the unobserved (latent) variables. In the case of LDA, the observed data are the words in the documents, and the latent variables are the topic assignments for each word and the topic distributions for each document.

The E-step of the EM algorithm for LDA involves computing the posterior distribution of the topic assignments given the observed words and the current estimates of the topic and document distributions. This posterior distribution is given by Bayes' rule:

$$p(z_{i,j} = k | w_{i,j}, \boldsymbol{\theta}_i, \boldsymbol{\phi}_k) = \frac{p(z_{i,j} = k | \boldsymbol{\theta}_i) p(w_{i,j} | z_{i,j} = k, \boldsymbol{\phi}_k)}{\sum_{k'=1}^{K} p(z_{i,j} = k' | \boldsymbol{\theta}_i) p(w_{i,j} | z_{i,j} = k', \boldsymbol{\phi}_{k'})} \tag{2.2}$$

where $z_{i,j}$ is the topic assigned to the jth word in the ith document, $w_{i,j}$ is the observed word, $\boldsymbol{\theta}_i$ is the topic distribution for the ith document, $\boldsymbol{\phi}_k$ is the word distribution for the kth topic, and K is the number of topics.

The posterior distribution gives the probability of each topic assignment for each word in each document, given the observed words and the current estimates of the topic and document distributions. These probabilities are used in the M-step of the EM algorithm to update the estimates of the topic and document distributions.

2.2.1.1.2 M-step The M-step [29] (Maximization step) of the EM (Expectation-Maximization) algorithm is the second step in each iteration of the algorithm. In this step, the algorithm updates the estimates of the model parameters to maximize the expected value of the complete-data log-likelihood function, computed in the E-step.

In the case of LDA, the M-step of the EM algorithm involves updating the estimates of the topic and document distributions based on the posterior distributions computed in the E-step. Specifically, the M-step involves maximizing the expected value of the complete-data log-likelihood function with respect to the topic and document distributions.

The update equations for the topic and document distributions in the M-step of the EM algorithm for LDA are as follows:

$$\phi_{k,w} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \mathbb{E}[z_{i,j} = k | w_{i,j}, \boldsymbol{\theta}_i, \boldsymbol{\phi}_k] \cdot \mathbb{I}(w_{i,j} = w)}{\sum_{w'=1}^{V} \sum_{i=1}^{M} \sum_{j=1}^{N} \mathbb{E}[z_{i,j} = k | w_{i,j}, \boldsymbol{\theta}_i, \boldsymbol{\phi}_k] \cdot \mathbb{I}(w_{i,j} = w')} \tag{2.3}$$

$$\theta_{i,k} = \frac{\sum_{j=1}^{N} \mathbb{E}[z_{i,j} = k | w_{i,j}, \boldsymbol{\theta}_i, \boldsymbol{\phi}_k]}{N_i} \tag{2.4}$$

where $\phi_{k,w}$ is the probability of word w in topic k, $\theta_{i,k}$ is the probability of topic k in document i, $z_{i,j}$ is the topic assigned to the jth word in the ith document, $w_{i,j}$ is the observed word, $\boldsymbol{\theta}_i$ is the topic distribution for the ith document, $\boldsymbol{\phi}_k$ is the word distribution for the kth topic, V is the size of the vocabulary, N_i is the number of words in the ith document, and $\mathbb{I}(w_{i,j} = w)$ is an indicator function that is 1 if $w_{i,j} = w$ and 0 otherwise.

The update equations for the topic and document distributions are derived by setting the derivatives of the expected complete-data log-likelihood function with respect to the topic and document distributions to zero and solving for the distributions.

After the topic and document distributions are updated in the M-step, the E-step is repeated to compute the posterior distributions of the topic assignments given the observed words and the updated estimates of the topic and document distributions. The algorithm iterates between the E-step and M-step until convergence.

2.2.1.2 Optimal number of topics

To obtain the optimal number of topics for a LDA model, a common approach is to create LDA models for different numbers of topics and then evaluate the models using coherence and Jaccard similarity measures. Coherence measures the degree of semantic similarity between high-scoring words in a topic, while Jaccard similarity measures the similarity between the sets of top words in different topics.

2.2.1.2.1 Coherence Coherence is a measure of the interpretability and semantic coherence of topics in a topic model [30]. It measures the degree of semantic similarity between high-scoring words in a topic. The coherence score for a topic is the average of the log probabilities of the co-occurrence frequencies of the top n words in the topic.

The co-occurrence frequency of two words w_i and w_j in a corpus is defined as the number of times they appear in the same context. The context of a word can be defined in various ways, such as the words that appear within a fixed window of the word in a document or the words that appear in the same sentence as the word.

The co-occurrence frequency of two words w_i and w_j in a corpus is denoted by $N(w_i, w_j)$. The probability of the co-occurrence of two words w_i and w_j in a corpus is given by:

$$P(w_i, w_j) = \frac{N(w_i, w_j)}{N} \tag{2.5}$$

where N is the total number of words in the corpus.

The pointwise mutual information (PMI) of two words w_i and w_j is a measure of their association that takes into account the frequency of the words in the corpus. The PMI of two words w_i and w_j is defined as:

$$\text{PMI}(w_i, w_j) = \log \frac{P(w_i, w_j)}{P(w_i)P(w_j)} \tag{2.6}$$

where $P(w_i)$ and $P(w_j)$ are the probabilities of the individual words in the corpus.

The coherence score for a topic is the average of the PMI scores of all pairs of words in the top n words of the topic. The coherence score for a topic is given by:

$$\text{Coherence}(T) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{PMI}(w_i, w_j) \tag{2.7}$$

where T is the topic, n is the number of words in the top n words of the topic, and w_i and w_j are the ith and jth words in the top n words of the topic.

The coherence score for the entire model is the average of the coherence scores for all topics in the model.

2.2.1.2.2 Jaccard similarity coefficient The Jaccard similarity (first presented as a ratio of verification in [31]) between two topics is computed as the size of the intersection of their top n words divided by the size of their union. The Jaccard similarity between all pairs of topics is computed, and the average Jaccard similarity for the model is the average of the Jaccard similarities for all pairs of topics.

Similarity of two sets U and V.

$$\text{Jaccard}(U, V) = \frac{|U \cap V|}{|U \cup V|} \tag{2.8}$$

where $|U|$ and $|V|$ are the sizes of the sets.

2.3 Criteria for evaluating the effectiveness of summarization system

2.3.1 General criteria

To evaluate the effectiveness of a summarization system, the following criteria can be considered:

- Accuracy: The summarization system should accurately reflect the content and essential points of the original proposal without distorting its meaning.

- Comprehensibility: The summary should be easy to understand for all members of the DAO, regardless of their technical expertise or linguistic background. The language used in the summary should be clear and concise, free from technical jargon or overly complex language.

- Relevance: The summary should focus on the most important points of the proposal, highlighting its key recommendations and the potential impact on the DAO. The summary should not include irrelevant information or details that do not contribute to the overall understanding of the proposal.

- Timeliness: The summarization system should be able to provide a summary of the proposal quickly and efficiently, allowing members to make informed decisions in a timely manner.

- Customization: The summarization system should be customizable to meet the specific needs and preferences of the DAO. For example, some members may prefer longer or more detailed summaries, while others may prefer shorter summaries that focus on the most critical points.

- Consistency: The summarization system should be consistent in its approach to summarizing proposals, using the same criteria and methods for all proposals to ensure fairness and impartiality.

By evaluating a summarization system based on these criteria, the system's effectiveness can be measured in terms of its ability to increase accessibility and participation in the decision-making process for DAO governance proposals.

2.3.2 ROUGE

ROUGE (Recall-Oriented Understudy for Gisting Evaluation [32]) is a set of metrics used to evaluate the quality of text summarization by comparing the generated summary to one or more reference summaries. ROUGE is widely used in scientific research for evaluating text summarization systems, including those based on machine learning.

$ROUGE$ measures the overlap between the generated summary and the reference summary in terms of n-gram co-occurrence. The most commonly used $ROUGE$ metrics are $ROUGE_N$ and $ROUGE_L$.

2.3.2.1 ROUGE-N

$ROUGE_N$ measures the overlap between the generated summary and the reference summary in terms of n-gram co-occurrence. The value of N can be any positive integer, with N=1 corresponding to unigrams, N=2 to bigrams, and so on. The $ROUGE_N$ score is calculated as follows:

$$ROUGE_N = \frac{\sum_{r \in R} \sum_{n \in grams(r,N)} Count_{match}(n, c)}{\sum_{r \in R} \sum_{n \in grams(r,N)} Count(n)} \tag{2.9}$$

where R is the set of reference summaries, c is the generated summary, and $Count(n)$ and $Count_{match}(n, c)$ are the number of occurrences of n in the reference summaries and the generated summary, respectively.

2.3.2.2 ROUGE-L

$ROUGE_L$ measures the longest common subsequence (LCS) between the generated summary and the reference summary. The $ROUGE_L$ score is calculated as follows:

$$ROUGE_L = \frac{\sum_{r \in R} \sum_{l \in LCS(r,c)} |l|}{\sum_{r \in R} |r|} \tag{2.10}$$

where $LCS(r, c)$ is the longest common subsequence between the reference summary r and the generated summary c.

$ROUGE$ metrics are widely used in scientific research for evaluating text summarization systems, including those based on machine learning. They provide a quantitative measure of the quality of the generated summary, which can be used to compare different summarization systems and to optimize their performance.

2.3.3 BERTScore

$BERTScore$ [33] is a metric used to evaluate the quality of text summarization by comparing the generated summary to one or more reference summaries. $BERTScore$ is based on the contextual embeddings generated by the BERT (Bidirectional Encoder Representations from Transformers) model.

$BERTScore$ measures the similarity between the generated summary and the reference summary in terms of contextual embeddings. The $BERTScore$ metric is calculated as follows:

$$\text{BERTScore}(P, R) = \frac{\sum_{i=1}^{n} F_{1,i} \cdot exp(-\frac{d_i^2}{2\sigma^2})}{\sum_{i=1}^{n} exp(-\frac{d_i^2}{2\sigma^2})} \tag{2.11}$$

where P is the set of predicted summaries, R is the set of reference summaries, n is the number of sentences in the summaries, $F_{1,i}$ is the harmonic mean of precision and recall for the i-th sentence, and d_i is the distance between the contextual embeddings of the i-th sentence in the predicted summary and the closest sentence in the reference summaries. σ is a smoothing parameter that controls the decay of the exponential function.

$BERTScore$ is also robust to variations in the length and content of the summaries, making it a reliable metric for evaluating text summarization systems.

2.3.4 BLEU

$BLEU$ (Bilingual Evaluation Understudy [34]) is a metric used to evaluate the quality of text summarization by comparing the generated summary to one or more reference summaries.

$BLEU$ measures the overlap between the generated summary and the reference summary in terms of n-gram co-occurrence. The most commonly used $BLEU$ metric is $BLEU_4$, which measures the overlap between the generated summary and the reference summary in terms of 4-gram co-occurrence. The $BLEU_4$ score is calculated as follows:

$$BLEU_4 = BP \cdot exp(\sum_{n=1}^{4} w_n log p_n) \tag{2.12}$$

where BP is the brevity penalty, which is a correction factor that penalizes summaries that are shorter than the reference summaries, w_n is the weight assigned to the n-gram precision, and p_n is the n-gram precision, which is the ratio of the number of n-grams in the generated summary that match an n-gram in the reference summary to the total number of n-grams in the generated summary.

The brevity penalty is calculated as follows:

$$BP = \begin{cases} 1 & \text{if } c > r \\ exp(1 - \frac{r}{c}) & \text{if } c \leq r \end{cases} \tag{2.13}$$

where c is the length of the generated summary and r is the length of the reference summary that has the closest length to the generated summary.

2.3.5 METEOR

METEOR (Metric for Evaluation of Translation with Explicit ORdering [35]) measures the similarity between the generated summary and the reference summary in terms of unigram matching, word order, and synonymy. The METEOR metric is calculated as follows:

$$METEOR = \frac{(\beta \cdot P + (1 - \beta) \cdot R)}{(\beta \cdot P + (1 - \beta) \cdot R + \gamma \cdot F)} \tag{2.14}$$

where P is the precision, R is the recall, F is the harmonic mean of precision and recall, β and γ are tunable parameters that control the relative importance of precision, recall, and F-score.

The precision is calculated as follows:

$$P = \frac{m}{m + a} \tag{2.15}$$

where m is the number of unigrams in the generated summary that match a unigram in the reference summary, and a is the number of unigrams in the generated summary that do not match any unigram in the reference summary.

The recall is calculated as follows:

$$R = \frac{m}{m + b} \tag{2.16}$$

where b is the number of unigrams in the reference summary that do not match any unigram in the generated summary.

METEOR also includes a penalty term for word order differences between the generated summary and the reference summary, which is calculated as follows:

$$\phi = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \delta_{i,j} \cdot \omega_{i,j}}{m} \tag{2.17}$$

where $\delta_{i,j}$ is the Kronecker delta function, which is 1 if the i-th unigram in the generated summary matches the j-th unigram in the reference summary, and 0 otherwise, $\omega_{i,j}$ is a weight assigned to the i-th unigram in the generated summary and the j-th unigram in the reference summary based on their position in the summary, and n is the length of the summary.

METEOR also includes a synonymy matching component, which is based on WordNet, a lexical database for English. The synonymy matching component is calculated as follows:

$$\psi = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} sim(w_i, w_j)}{m} \tag{2.18}$$

where $sim(w_i, w_j)$ is the maximum similarity score between the i-th unigram in the generated summary and the j-th unigram in the reference summary based on their synonyms in WordNet.

2.3.6 GRAD

The GRAD (GRAph Distance [36]) metric is a graph-based measure that aims to estimate how well summary terms are connected to full text terms. It is based on the assumption that a good summary is made of the terms that refer to the central vertices in the semantic graph, i.e. the terms that are connected to the maximal number of other terms in a full text. According to this metric, the score of a summary is estimated as a normalized inverted sum of distances from every term in the text to its closest term appearing in the summary S:

$$score_S = \frac{1}{\sum_{v \in V} \min_{v_j \in S} d(v, v_j)} \tag{2.19}$$

where $d(v_j, v_i)$ is the shortest path between v_i and v_j. To calculate minimal distances from every term in the text to its closest term from the summary, the authors used the Dijkstra algorithm. The obtained results show that the GRAD measure significantly outperformed overlap-based baselines on both test collections in distinguishing human written abstracts from generated summaries of poor quality.

2.3.7 Pyramid

Pyramid [37] is a human-based evaluation method that compares the system-generated summaries with the reference summaries created by humans. Pyramid metric is based on the idea that a good summary should cover all the important information in the source text and should not include any irrelevant information.

The Pyramid metric consists of three components: coverage, density, and diversity.

Coverage measures the percentage of important information in the source text that is covered by the summary. It is calculated as follows:

$$Coverage = \frac{\sum_{s \in S} \sum_{w \in s} \min(count_{ref}(w), count_{sys}(w))}{\sum_{s \in S} \sum_{w \in s} count_{ref}(w)} \tag{2.20}$$

where S is the set of sentences in the source text, $count_{ref}(w)$ is the number of times the word w appears in the reference summary, and $count_{sys}(w)$ is the number of times the word w appears in the system-generated summary.

Density measures the degree of redundancy in the summary. It is calculated as follows:

$$Density = \frac{\sum_{s \in S} \sum_{w \in s} \min(count_{sys}(w), count_{sys}(w))}{\sum_{s \in S} \sum_{w \in s} count_{sys}(w)} \tag{2.21}$$

where $count_{sys}(w)$ is the number of times the word w appears in the system-generated summary.

Diversity measures the degree to which the summary contains different types of information. It is calculated as follows:

$$Diversity = 1 - \frac{\sum_{s \in S} \max_{r \in R}(\text{sim}(s, r))}{\sum_{s \in S} \sum_{r \in R} \text{sim}(s, r)} \tag{2.22}$$

where R is the set of reference summaries, $\text{sim}(s, r)$ is the cosine similarity between sentence s in the system-generated summary and sentence r in the reference summary.

The final Pyramid score is calculated as the geometric mean of the three components:

$$Pyramid = \sqrt{Coverage \times Density \times Diversity} \tag{2.23}$$

The Pyramid metric has been shown to be a reliable and consistent evaluation method for text summarization systems.

2.3.8 NIST

The NIST (National Institute of Standards and Technology [38]) metric is another widely used method for evaluating text summarization systems. It is a variant of the ROUGE (Recall-Oriented Understudy for Gisting Evaluation) metric, which measures the overlap between the system-generated summaries and the reference summaries created by humans.

The NIST metric is based on the idea that a good summary should contain important information from the source text and should not include any irrelevant information. It consists of two components: unigram precision and unigram recall.

Unigram Precision measures the percentage of unigrams (single words) in the system-generated summary that also appear in the reference summaries. It is calculated as follows:

$$UnigramPrecision = \frac{\sum_{w \in s} \min(count_{ref}(w), count_{sys}(w))}{\sum_{w \in s} count_{sys}(w)} \tag{2.24}$$

where $count_{ref}(w)$ is the number of times the word w appears in the reference summaries, and $count_{sys}(w)$ is the number of times the word w appears in the system-generated summary.

Unigram Recall measures the percentage of unigrams in the reference summaries that also appear in the system-generated summary. It is calculated as follows:

$$UnigramRecall = \frac{\sum_{w \in s} \min(count_{ref}(w), count_{sys}(w))}{\sum_{w \in s} count_{ref}(w)} \tag{2.25}$$

The final NIST score is calculated as the geometric mean of the unigram precision and unigram recall:

$$NIST = \sqrt{UnigramPrecision \times UnigramRecall} \tag{2.26}$$

The NIST metric has been shown to be a reliable and consistent evaluation method for text summarization systems.

2.4 Summarization approaches

2.4.1 Types of summarization

There are two main approaches to summarization: abstractive and extractive summarization.

Abstractive Summarization [39] involves generating a summary that captures the essential meaning of the text using natural language generation techniques. This approach requires the system to understand the content of the text and generate a new summary that may not contain the exact phrases or sentences from the original text. Abstractive summarization can generate more concise and readable summaries, but it requires a deep understanding of the text, which is challenging to achieve with current machine learning techniques.

Extractive summarization [40] involves selecting the most critical sentences or phrases from the text to form a summary. This approach does not generate new sentences or phrases, but instead extracts the most relevant information from the text to form a summary. Extractive summarization is often more straightforward to implement and can generate more accurate summaries, but may be less readable or concise than abstractive summarization.

In the context of DAO governance proposals, both abstractive and extractive summarization approaches can be used. Abstractive summarization can be useful for capturing the essential meaning of a proposal and generating a concise and readable summary. Extractive summarization can be useful for highlighting the most critical points and recommendations in a proposal, which can help readers quickly understand its significance. The choice of summarization approach may depend on the specific needs and preferences of the DAO members and the nature of the proposal being summarized.

2.4.2 Advantages and disadvantages

In the context of DAO governance proposals, each approach has its advantages and disadvantages.

Abstractive summarization involves generating a summary by understanding the meaning of the text and using natural language processing techniques to generate a new summary in the same language. One advantage of abstractive summarization is that it can generate more concise summaries than extractive summarization, as it is not limited to selecting and reordering sentences from the original text. Another advantage is that it can capture the essential meaning of the text more accurately and can convey the author's intended message more effectively. However, a disadvantage of abstractive summarization is that it can be computationally expensive and requires a large amount of training data to generate high-quality summaries. Additionally, because it generates a new summary, it can introduce errors or inaccuracies if the generated summary does not reflect the original text's meaning accurately.

Extractive summarization involves selecting and rearranging sentences from the original text to create a summary that captures the essential information. One advantage of extractive summarization is that it is less computationally expensive and requires less training data than abstractive summarization. Additionally, because it directly extracts sentences from the original text, it is less likely to introduce errors or inaccuracies. However, a disadvantage of extractive summarization is that it can produce summaries that are longer and less concise than abstractive summarization. Additionally, it can struggle to capture the author's intended message when the original text is long and complex, and it may not capture the most critical information in the text.

Overall, both abstractive and extractive summarization approaches have advantages and disadvantages in the context of DAO governance proposals. The choice of approach may also depend on the specific needs of the DAO and the nature of the governance proposal being summarized.

2.5 SOTA models overview

2.5.1 GPT

2.3

Models	Speed + Availability	Quality of results	Fine Tuning	Price input [1K tokens*]	Price output [1K tokens*]	Capacity [tokens]
GPT-3 (Davinci)	✓	✓	✓	$0.12	$0.12	4097
ChatGPT	✓	✓	✗	$0.0015	$0.002	4096
GPT-4	✗	✓	✗	$0.03	$0.06	8192
GPT-4-32k	✗	✓	✗	$0.06	$0.12	32768

Figure 2.3: GPT models comparision

2.5.1.1 Fine-tuning

2.5.1.2 GPT3.5

2.5.1.3 GPT4

2.5.2 BART

Chapter 3

Design

3.1 System design

3.1

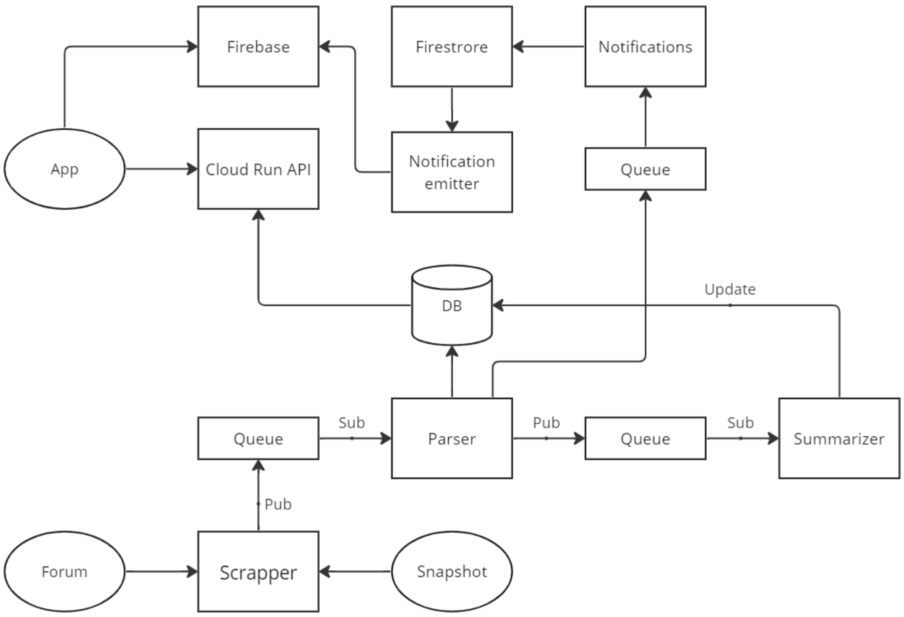


Figure 3.1: Backend

3.2 Frontend

3.2

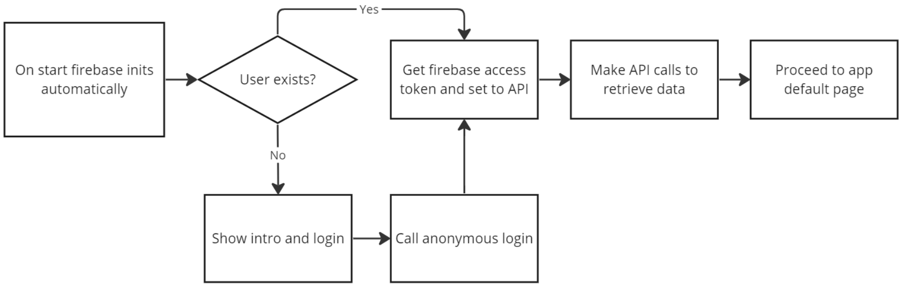


Figure 3.2: Frontend

3.3 UI/UX design

3.3

3.4 Application

3.4

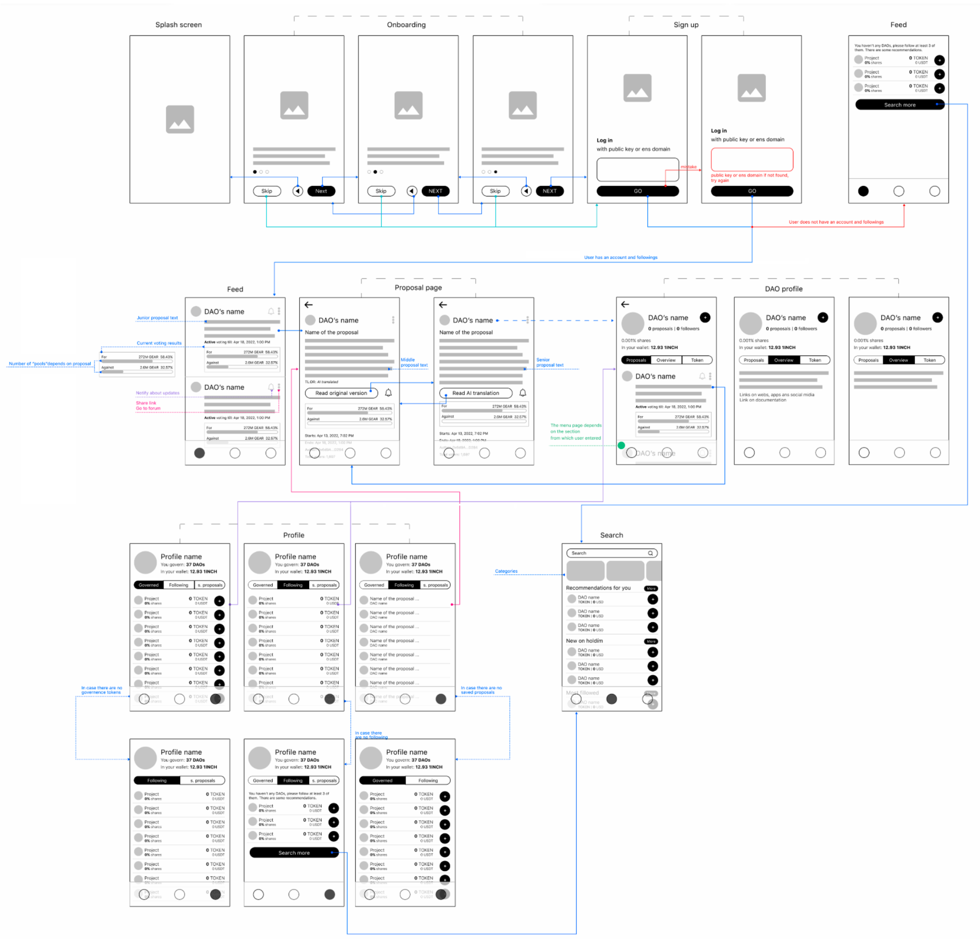
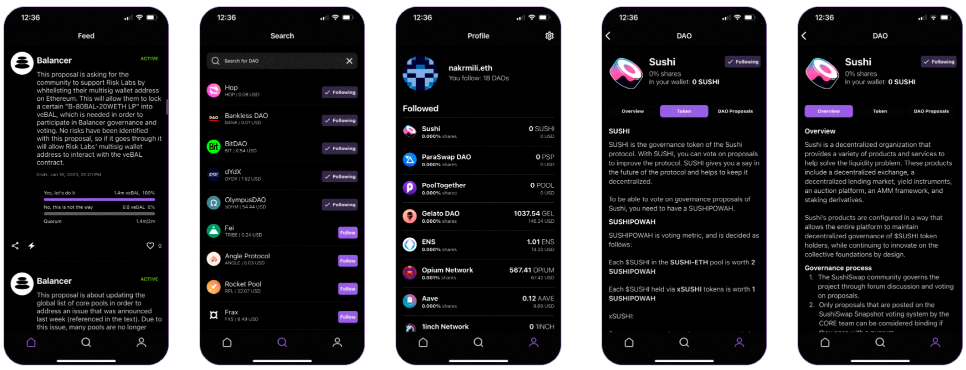


Figure 3.3: UI/UX design



(a) Generated summaries governance proposals feed

(b) Interface to follow DAOs

(c) User profile

(d) Information about tokens

(e) Overview on DAO processes and procedures

Figure 3.4: Public beta application design

Chapter 4

Results and discussion

4.1 Dataset collection

4.2 EDA results

4.2.1 Wordclouds

4.1

(a) Proposals texts

(b) Summaries texts

Figure 4.1: Wordclouds

4.2.2 Analyze relevant feature distributions

4.2.2.1 Source documents data distributions

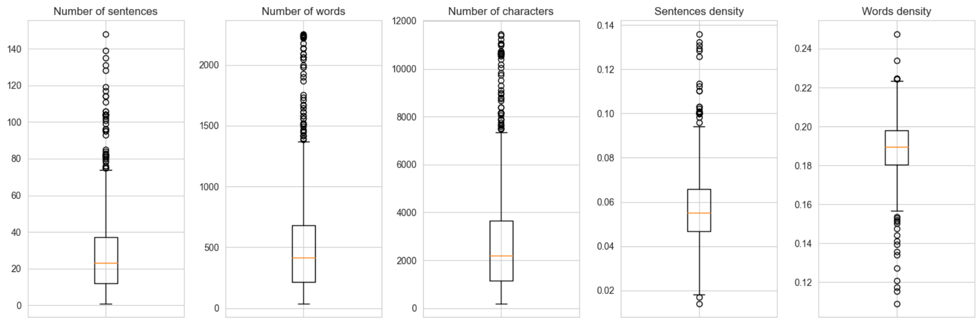
4.2

4.2.2.2 Ground truth summaries data distributions

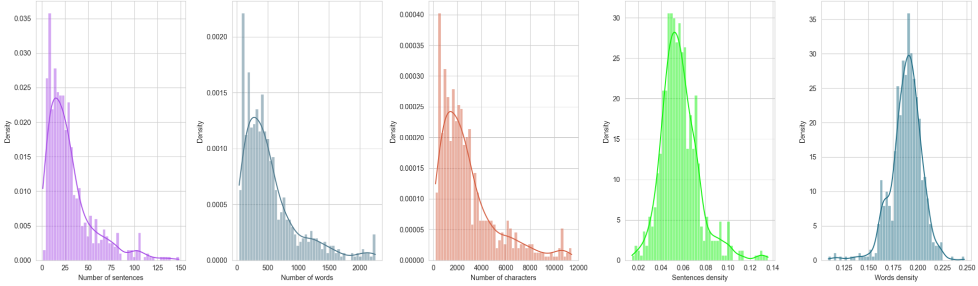
4.3

4.2.2.3 Categorizing and POS tagging words

4.4



(a) Features box plots



(b) Features histograms

Figure 4.2: Source data distributions

4.2.2.4 Use of stopwords and punctuations

4.5

4.2.3 Topic modeling

4.2.3.1 Optimal Number of Topics for LDA

To apply LDA to a collection of source texts, the first step is to preprocess the texts by tokenizing them, removing stop words, and stemming the remaining words. Then, the LDA model is trained on the preprocessed texts using a library such as Gensim [41] in Python [42]. The number of topics is a hyperparameter that must be chosen by the user. After training, the model can be used to transform new documents into topic distributions.

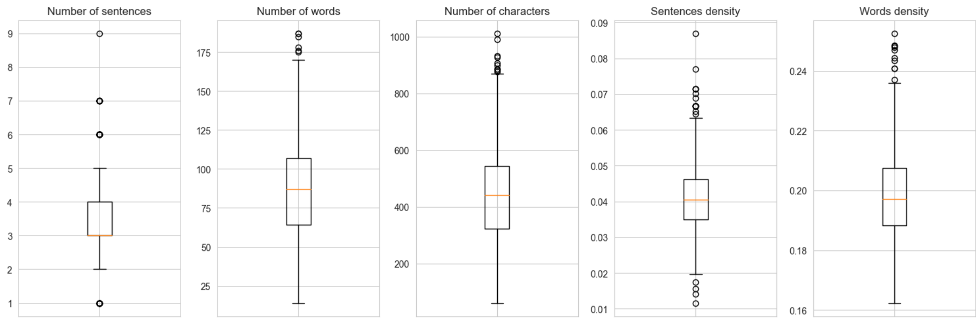
4.6

4.2.3.2 Visualization

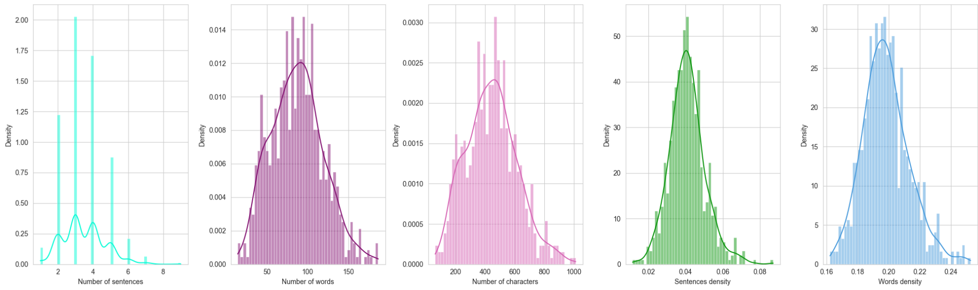
4.7

4.2.3.3 Topic analysis

4.1



(a) Features box plots



(b) Features histograms

Figure 4.3: Ground truth summaries data distributions

4.3 Generating summaries

4.3.1 GPT finetuning

4.8

4.4 Algorithms evaluation

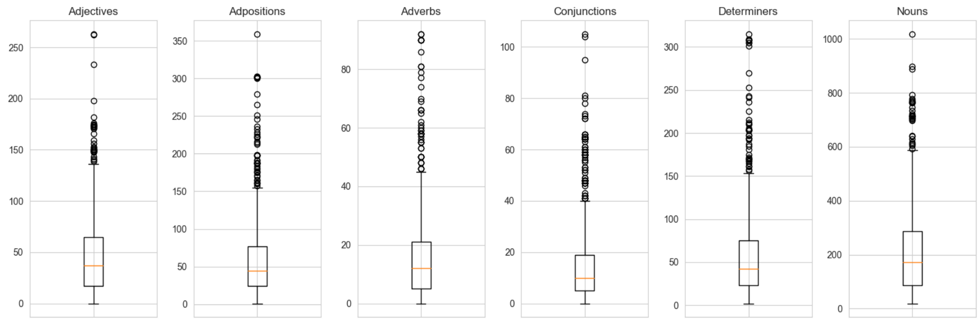
4.4.1 ROUGE, ROUGE aggregated, Bert score, BLEU, METEOR

4.9

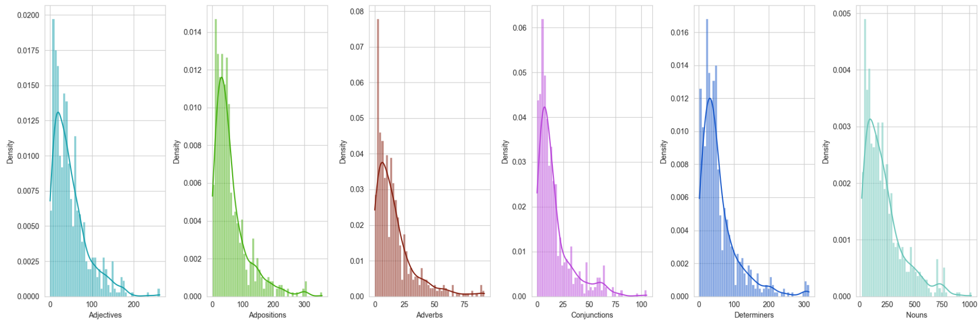
4.10

4.4.2 G-eval inspired metric

Here we develop a demonstration of a reference-free text evaluation system utilizing GPT-4, drawing inspiration from the G-Eval TODO (add source) framework, which gauges the excellence of generated text employing extensive language models. Diverging from conventional metrics such as ROUGE or BERTScore, which necessitate reference summaries for comparison, GPT-4-based evaluator evaluates the caliber of generated content exclusively based on



(a) Features box plots



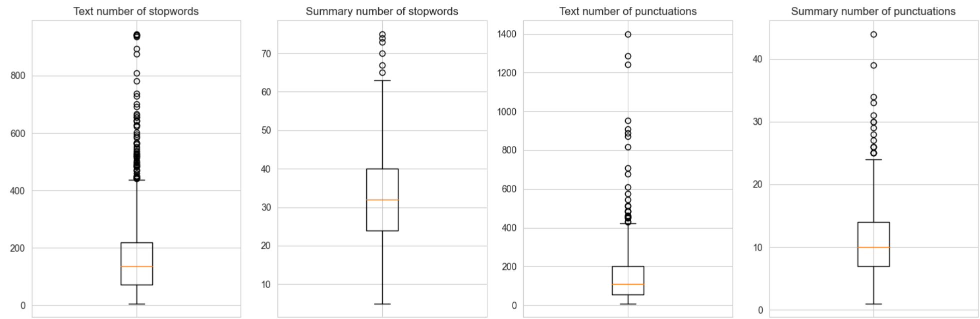
(b) Features histograms

Figure 4.4: POS tagging words distributions

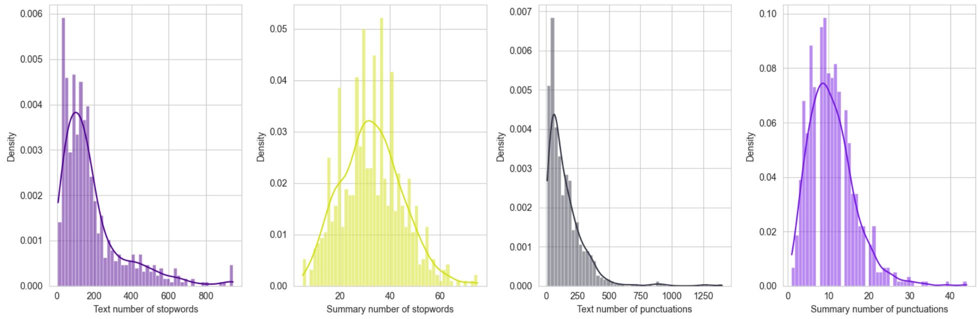
the input prompt and text, without relying on any ground truth references. This characteristic renders it suitable for application in novel datasets and tasks characterized by sparse or inaccessible human references.

In this study, we present an outline of our approach as follows:

1. We establish four distinct criteria:
    - **Relevance:** Evaluates whether the summary encompasses solely crucial information while eliminating redundancies.
    - **Coherence:** Assesses the logical progression and organization of the summary.
    - **Consistency:** Verifies the alignment of the summary with the factual content of the source document.
    - **Fluency:** Rates the grammatical correctness and readability of the summary.
2. For each of these criteria, we design specific prompts. These prompts are formulated considering both the original document and the summary as inputs. Leveraging chain-of-thought generation techniques, we guide the model to produce a numerical score ranging from 1 to 5 for each criterion.



(a) Features box plots



(b) Features histograms

Figure 4.5: Stopwords and punctuation distributions

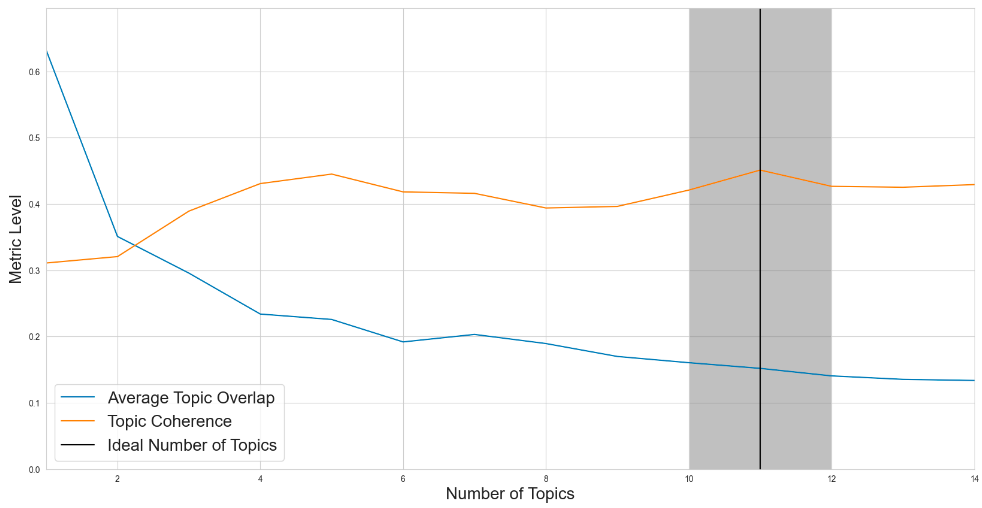


Figure 4.6: Optimal number of topics metrics for LDA model

3. Using GPT-4, we generate scores based on the defined prompts and apply them across multiple summaries for comparison.

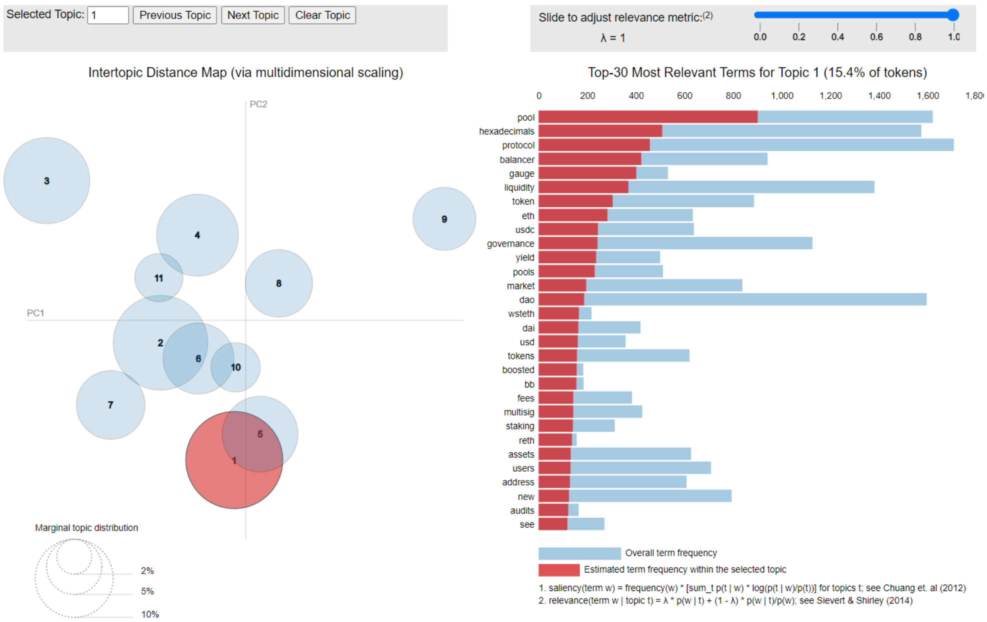


Figure 4.7: Topic Modeling with Gensim (Python)

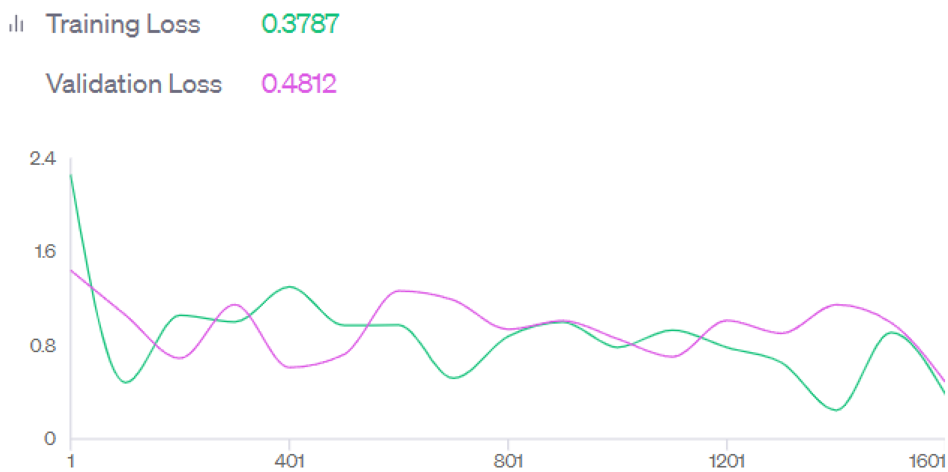


Figure 4.8: Finetuning process

4.4.2.1 Evaluation prompt template

4.1

Topic label	Highly associated terms	ID
Liquidity Pool Governance	pool, hexadecimals, protocol, balancer, gauge, liquidity, token, eth, usdc, governance	1
Ethereum Ecosystem Insights	inch, chain, uniswap, network, ethereum, protocol, security, governance, community, token	2
Community Governance Proposals	dao, proposal, governance, community, vote, working, grants, funding, process, group	3
Frax Token Bridging Strategies	frax, ethereum, fraxferry, tokens, token, deploy, crv, address, bridging, frxeth	4
Stablecoin Protocol Strategies	protocol, ageur, liquidity, aggold, pool, proposal, chain, angle, paxg, stablecoin	5
Balancer Liquidity Management	balancer, bal, liquidity, yield, stargate, protocol, vebal, pool, dao, weth	6
Arbitrum Governance Decisions	hexadecimals, address, dao, uint, pool, multisig, proposal, bytes, contract, arbitrum	7
Interest Rate Proposal	rate, interest, reserve, proposal, slope, bal, liquidity, balancer, borrowing, curve	8
Risk and Liquidity Management	risk, aave, liquidation, parameter, liquidity, polygon, parameters, market, supply, value	9
Frax Token Deployment Focus	frax, ethereum, fraxferry, tokens, token, deploy, crv, address, bridging, frxeth	10
Market Governance Options	proposal, dao, market, ghst, treasury, liquidity, community, option, price, protocol	11

Table 4.1: Named topic labels

```
EVALUATION_PROMPT_TEMPLATE = """
You will be given summaries written for an article. Your task is to rate the summary on four metrics: Relevance, Coherence, Consistency, and Fluency.
Please make sure you read and understand these instructions very carefully.
Please keep this document open while reviewing, and refer to it as needed.

Evaluation Criteria:{criteria1}
Evaluation Steps:{steps1}
Evaluation Criteria:{criteria2}
Evaluation Steps:{steps2}
Evaluation Criteria:{criteria3}
Evaluation Steps:{steps3}
Evaluation Criteria:{criteria4}
Evaluation Steps:{steps4}
Example:
Source Text:
{document}
Summary 1:
{summary1}
Summary 2:
{summary2}
Summary 3:
```

ROUGE

4 rows × 4 columns pd.DataFrame

Model	rouge1	rouge2	rougeL	rougeLsum
GPT3_5_turbo	[0.36464…	[0.17222…	[0.27071…	[0.27071823…
finetuned_GPT3_5_turbo	[0.43544…	[0.22391…	[0.32911…	[0.32911392…
GPT4	[0.42519…	[0.18997…	[0.30446…	[0.30446194…
BART_large_CNN	[0.40350…	[0.38235…	[0.33333…	[0.33333333…

ROUGE aggregated

4 rows × 4 columns pd.DataFrame

Model	rouge1	rouge2	rougeL	rougeLsum
GPT3_5_turbo	0.329750	0.163066	0.237174	0.237766
finetuned_GPT3_5_turbo	0.279492	0.147078	0.191631	0.190559
GPT4	0.297198	0.111349	0.197919	0.197961
BART_large_CNN	0.444843	0.414788	0.422422	0.422622

Bert score

4 rows × 3 columns pd.DataFrame

Model	precision	recall	f1
GPT3_5_turbo	[0.86074686…	[0.9066402…	[0.883097…
finetuned_GPT3_5_turbo	[0.87143838…	[0.9032419…	[0.887055…
GPT4	[0.86171454…	[0.9043322…	[0.882509…
BART_large_CNN	[0.86292219…	[0.9446718…	[0.901948…

Figure 4.9: ROUGE, ROUGE aggregated, Bert score

Score criteria and steps are as follows.

4.4.2.2 Metric 1: Relevance

Relevancy Score Criteria:
Relevance (1-5) - selection of important content from the source. The summary should include only important information from the source document.

BLEU

4 rows × 6 columns pd.DataFrame

Model	bleu	precisions	brevity_penalty	length_ratio	translation_length	reference_length
GPT3_5_turbo	0.049632	[0.128468757…	1.0	6.095476	89189	14632
finetuned_GPT3_5_turbo	0.048647	[0.102804157…	1.0	7.726674	89189	11543
GPT4	0.032853	[0.120160557…	1.0	5.661356	89189	15754
BART_large_CNN	0.231724	[0.256859029…	1.0	3.809705	89189	23411

METEOR

4 rows × 1 columns pd.DataFrame

	meteor
GPT3_5_turbo	0.396429
finetuned_GPT3_5_turbo	0.375044
GPT4	0.337844
BART_large_CNN	0.617220

Figure 4.10: BLEU, METEOR

Annotators were instructed to penalize summaries which contained redundancies and excess information.

Relevancy Score Steps:

1. Read the summary and the source document carefully.
2. Compare the summary to the source document and identify the main points of the article.
3. Assess how well the summary covers the main points of the article, and how much irrelevant or redundant information it contains.
4. Assign a relevance score from 1 to 5.

4.4.2.3 Metric 2: Coherence

Coherence Score Criteria:

Coherence (1-5) - the collective quality of all sentences. We align this dimension with the DUC quality question of structure and coherence whereby "the summary should be well-structured and well-organized. The summary should not just be a heap of related information, but should build from sentence to a coherent body of information about a topic."

Coherence Score Steps:

1. Read the article carefully and identify the main topic and key points.
2. Read the summary and compare it to the article. Check if the summary covers the main topic and key points of the article, and if it presents them in a clear and logical order.
3. Assign a score for coherence on a scale of 1 to 5, where 1 is the lowest and 5 is the highest based on the Evaluation Criteria.

4.4.2.4 Metric 3: Consistency

Consistency Score Criteria:
Consistency (1-5) - the factual alignment between the summary and the summarized source. A factually consistent summary contains only statements that are entailed by the source document. Annotators were also asked to penalize summaries that contained hallucinated facts.

Consistency Score Steps:

1. Read the article carefully and identify the main facts and details it presents.

2. Read the summary and compare it to the article. Check if the summary contains any factual errors that are not supported by the article.

3. Assign a score for consistency based on the Evaluation Criteria.

4.4.2.5 Metric 4: Fluency

Fluency Score Criteria:
Fluency (1-3): the quality of the summary in terms of grammar, spelling, punctuation, word choice, and sentence structure. 1: Poor. The summary has many errors that make it hard to understand or sound unnatural. 2: Fair. The summary has some errors that affect the clarity or smoothness of the text, but the main points are still comprehensible. 3: Good. The summary has few or no errors and is easy to read and follow.

Fluency Score Steps:
Read the summary and evaluate its fluency based on the given criteria. Assign a fluency score from 1 to 3.

4.4.2.6 Results

4.11

1-10 121 rows × 17 columns pd.DataFrame

	seniorDescription	sum1_Relevance	sum1_Coherence	sum1_Consistency	sum1_Fluency	sum2_Relevance
0	title: Aave V3 Harm…	4	5	5	3	4
1	This proposals expe…	4	4	4	3	5
2	Authors: [Noodles](…	4	4	4	3	4
3	Summary: Enable a 1…	4	4	4	3	5
4	This vote proposes …	4	5	5	3	5
5	Summary. The Meta-G…	4	5	5	3	5
6	This proposals expe…	4	5	5	3	4
7	Description. Propos…	4	4	4	3	5
8	[Read the full prop…	4	4	4	3	5

Figure 4.11: Custom metric inspired by G-Eval

Conclusion and future work

In conclusion, this thesis has addressed the pressing need for effective text summarization techniques within Decentralized Autonomous Organizations (DAOs) to enhance accessibility and participation in decision-making processes. Through a comprehensive exploration of DAOs, their governance structures, and the challenges associated with summarizing governance proposals, this research has contributed valuable insights to the field.

The development and evaluation of a personalized machine learning-based summarization system, tailored specifically for DAO governance proposals, have demonstrated promising results. The system, leveraging state-of-the-art models such as GPT-4, has shown improvements in accessibility and participation by providing concise and understandable summaries of complex governance proposals.

The evaluation of the system using metrics such as accuracy, comprehensibility, and relevance has highlighted its effectiveness in addressing the identified challenges. By comparing the results with existing approaches, this thesis has underscored the importance of customized summarization systems in facilitating informed decision-making within DAOs.

Looking ahead, further research is warranted to explore additional refinements and enhancements to the summarization system. Future studies could investigate the integration of more advanced natural language processing techniques, as well as the incorporation of user feedback mechanisms to continuously improve the system's performance.

Overall, this thesis serves as a foundation for future research endeavors aimed at advancing text summarization methodologies within the context of DAO governance. By enabling greater accessibility and participation, these efforts contribute to the evolution and democratization of decision-making processes within decentralized ecosystems.

Bibliography

[1] Kondova, G.; Barba, R. Governance of decentralized autonomous organizations. *Journal of Modern Accounting and Auditing*, volume 15, no. 8, 2019: pp. 406–411.

[2] Pradha, S.; Halgamuge, M. N.; et al. Effective text data preprocessing technique for sentiment analysis in social media data. In *2019 11th international conference on knowledge and systems engineering (KSE)*, IEEE, 2019, pp. 1–8.

[3] Rikken, O.; Janssen, M.; et al. The ins and outs of decentralized autonomous organizations (DAOs) unraveling the definitions, characteristics, and emerging developments of DAOs. *Blockchain: Research and Applications*, 2023: p. 100143.

[4] Wang, F.-Y.; Ding, W.; et al. The DAO to DeSci: AI for free, fair, and responsibility sensitive sciences. *IEEE Intelligent Systems*, volume 37, no. 2, 2022: pp. 16–22.

[5] Bellavitis, C.; Fisch, C.; et al. The rise of decentralized autonomous organizations (DAOs): a first empirical glimpse. *Venture Capital*, volume 25, no. 2, 2023: pp. 187–203.

[6] Nguyen, T.-V.; Tran, A.-T.; et al. Information fusion on delivery: A survey on the roles of mobile edge caching systems. *Information Fusion*, volume 89, 2023: pp. 486–509.

[7] Axelsen, H.; Jensen, J. R.; et al. When is a DAO Decentralized? *arXiv preprint arXiv:2304.08160*, 2023.

[8] Wu, D.; Yang, Z.; et al. Virtual-Reality Inter-Promotion Technology for Metaverse: A Survey. *IEEE Internet of Things Journal*, 2023.

[9] Dao, T. M.; McGroarty, F.; et al. The Brexit vote and currency markets. *Journal of International Financial Markets, Institutions and Money*, volume 59, 2019: pp. 153–164.

[10] Johansen, B.; Johansen, R.; et al. *Office Shock: Creating Better Futures for Working and Living*. Berrett-Koehler Publishers, 2023.

[11] Augustin, N.; Eckhardt, A.; et al. Understanding decentralized autonomous organizations from the inside. *Electronic Markets*, volume 33, no. 1, 2023: pp. 1–14.

[12] Forum, W. E. Decentralized Autonomous Organization Toolkit. *Insight Report*, 2023, in collaboration with the Wharton Blockchain and Digital Asset Project.

[13] Lee, C. A.; Chow, K.; et al. Decentralized governance and artificial intelligence policy with blockchain-based voting in federated learning. *Frontiers in Research Metrics and Analytics*, volume 8, 2023: p. 103.

[14] Reiff, N. Decentralized autonomous organization (DAO): Definition, purpose, and example. 2022.

[15] Yu, G.; Wang, Q.; et al. Leveraging Architectural Approaches in Web3 Applications-A DAO Perspective Focused. In *2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, IEEE, 2023, pp. 1–6.

[16] Wang, Q.; Li, R.; et al. Non-fungible token (NFT): Overview, evaluation, opportunities and challenges. *arXiv preprint arXiv:2105.07447*, 2021.

[17] Sims, A. Decentralised autonomous organisations: governance, dispute resolution and regulation. *Dispute Resolution and Regulation (May 31, 2021)*, 2021.

[18] Venugopalan, S.; Aydt, H. Dance of the DAOs: Building Data Assets as a Use Case. *arXiv preprint arXiv:2301.05945*, 2023.

[19] Saito, Y.; Rose, J. Reputation-based decentralized autonomous organization for the non-profit sector: leveraging blockchain to enhance good governance. *Frontiers in Blockchain*, volume 5, 2023, doi:10.3389/fbloc.2022.1083647.

[20] Hsieh, Y.-Y. *The rise of decentralized autonomous organizations: Coordination and growth within cryptocurrencies*. Dissertation thesis, The University of Western Ontario (Canada), 2018.

[21] Snapshot. `https://snapshot.org/`, accessed: October 9, 2023.

[22] GraphQL API. https://www.freecodecamp.org/news/building-consuming-and-documenting-a-graphql-api/, accessed: May 2, 2023.

[23] Batista, L.; Alexandre, L. A. Text Pre-processing for Lossless Compression. In *DCC*, Citeseer, 2008, p. 506.

[24] Savaliya, B. R.; Philip, C. G. Email fraud detection by identifying email sender. In *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, IEEE, 2017, pp. 1420–1422.

[25] Vining, G.; Judd, C.; et al. Data analysis: a model-comparison approach. *Technometrics*, volume 34, 1992: p. 107, doi:10.2307/1269570.

[26] Rasool, A.; Bunterngchit, C.; et al. Improved machine learning-based predictive models for breast cancer diagnosis. *International Journal of Environmental Research and Public Health*, volume 19, 2022: p. 3211, doi:10.3390/ijerph19063211.

[27] Conz, E.; Denicolai, S.; et al. The resilience strategies of smes in mature clusters. *Journal of Enterprising Communities People and Places in the Global Economy*, volume 11, 2017: pp. 186–210, doi:10.1108/jec-02-2015-0015.

[28] Ahmed, M. Intelligent big data summarization for rare anomaly detection. *Ieee Access*, volume 7, 2019: pp. 68669–68677, doi:10.1109/access.2019.2918364.

[29] Blei, D. M.; Ng, A. Y.; et al. Latent dirichlet allocation. *Journal of machine Learning research*, volume 3, no. Jan, 2003: pp. 993–1022.

[30] Röder, M.; Both, A.; et al. Exploring the space of topic coherence measures. In *Proceedings of the eighth ACM international conference on Web search and data mining*, 2015, pp. 399–408.

[31] Murphy, A. H. The Finley affair: A signal event in the history of forecast verification. *Weather and forecasting*, volume 11, no. 1, 1996: pp. 3–20.

[32] Lin, C.-Y. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, 2004, pp. 74–81.

[33] Zhang, T.; Kishore, V.; et al. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*, 2019.

[34] Papineni, K.; Roukos, S.; et al. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 2002, pp. 311–318.

[35] Banerjee, S.; Lavie, A. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In *Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization*, 2005, pp. 65–72.

[36] Ermakova, L.; Firsov, A. GRAD: A Metric for Evaluating Summaries. In *Coria*, 2018.

[37] Nenkova, A.; Passonneau, R. J. Evaluating content selection in summarization: The pyramid method. In *Proceedings of the human language technology conference of the north american chapter of the association for computational linguistics: Hlt-naacl 2004*, 2004, pp. 145–152.

[38] Przybocki, M.; Peterson, K.; et al. The NIST 2008 Metrics for machine translation challenge—overview, methodology, metrics, and results. *Machine Translation*, volume 23, 2009: pp. 71–103.

[39] Nallapati, R.; Zhou, B.; et al. Abstractive text summarization using sequence-to-sequence rnns and beyond. *arXiv preprint arXiv:1602.06023*, 2016.

[40] Pardo, T. A. S.; Rino, L. H. M.; et al. Extractive summarization: how to identify the gist of a text. In *Proceedings of I2TS'2002-International Information Technology Symposium*, 2002.

[41] Řehůřek, R. Software Framework for Topic Modelling with Large Corpora. LREC 2010, 2010. Available from: https://radimrehurek.com/gensim/

[42] Foundation, P. S. Python. 1991–. Available from: https://www.python.org/

Appendix **A**

Acronyms

DAO Decentralized autonomous organization

EDA Exploratory data analysis

NLP Natural Language Processing

GPT Generative Pre-trained Transformer

UI/UX User Interface/User Experience

LDA Latent Dirichlet Allocation

ROUGE Recall-Oriented Understudy for Gisting Evaluation

BLEU Bilingual Evaluation Understudy

METEOR Metric for Evaluation of Translation with Explicit ORdering

GRAD Glossing and Rating of Automated Summarization and Translation

NIST National Institute of Standards and Technology

Appendix **B**

Contents of enclosed CD

```
data_analysis ........................ directory for data analysis scripts
├── models ......................... directory for storing info about models
├── data .............................................. directory for data
├── metrics ............................... directory for evaluation metrics
├── notebooks ........................... directory for Jupyter notebooks
└── README.md ............................... file with project description
```