Chatbots For Social Change

March 11, 2024

On the 28th of April 2012 the contents of the English as well as German Wikibooks and Wikipedia projects were licensed under Creative Commons Attribution-ShareAlike 3.0 Unported license. A URI to this license is given in the list of figures on page 61. If this document is a derived work from the contents of one of these projects and the content was still licensed by the project under this license at the time of derivation this document has to be licensed under the same, a similar or a compatible license, as stated in section 4b of the license. The list of contributors is included in chapter Contributors on page 59. The licenses GPL, LGPL and GFDL are included in chapter Licenses on page 65, since this book and/or parts of it may or may not be licensed under one or more of these licenses, and thus require inclusion of these licenses. The licenses of the figures are given in the list of figures on page 61. This PDF was generated by the LATEX typesetting software. The LATEX source code is included as an attachment (`source.7z.txt`) in this PDF file. To extract the source from the PDF file, you can use the `pdfdetach` tool including in the `poppler` suite, or the `http://www.pdflabs.com/tools/pdftk-the-pdf-toolkit/` utility. Some PDF viewers may also let you save the attachment to a file. After extracting it from the PDF file you have to rename it to `source.7z`. To uncompress the resulting archive we recommend the use of `http://www.7-zip.org/`. The LATEX source itself was generated by a program written by Dirk Hünniger, which is freely available under an open source license from `http://de.wikibooks.org/wiki/Benutzer:Dirk_Huenniger/wb2pdf`.

Contents

1 Introduction

By necessity, this book is widely interdisciplinary, bringing together insights from scholarly work understanding "understanding," social action, social systems, the social psychology of belief, the philosophy of science, the sociology of belief systems, research ethics, ethics of privacy, and of interaction, clinical psychology, the technical intricacies of LLMs, frameworks of knowledge management, automated proof-checking, to name some of the most important fields of knowledge involved.

As you can see, this textbook cannot be built by just one person. I (Alec McGail) am writing this now to start the endeavor in a free and transparent manner, very much in the spirit of what's discussed in *Section 2: What's Ethical?*. So, anyone who feels they have something they can contribute to the endeavor should reach out to me (am2873@cornell.edu), or just go ahead and make changes.

If you'd like to follow my process in working through this class, follow my Twitch channel[1] and YouTube channel[2].

Here, you will embark on an intellectual adventure, blending the theoretical intricacies of intersubjective thought with hands-on training in Large Language Models (LLMs). By the end, you won't just understand the mechanics of these digital marvels; you will be the craftsman behind their creation.

For the intrepid scholar and the visionary educator alike, this journey promises a harmonious blend of robust theoretical foundations and cutting-edge practical applications. Each week unfurls a new layer of understanding, from ethical considerations to technical mastery, all culminating in a capstone project where you breathe life into your very own chatbot. This isn't just a course; it's a call to be at the forefront of a sociotechnological revolution, with the power to shape discourse, challenge beliefs, and unite our ever-evolving global community. Get ready to be both the student and the pioneer, charting the path for the next wave of societal evolution.

Disclaimer: *"Chatbots for Social Change"* has been collaboratively developed with the aid of ChatGPT, a product of OpenAI's cutting-edge Large Language Model (LLM) technology. The utilization of ChatGPT in the creation of this WikiBook is a practical demonstration of the subject matter at the heart of this course. As learners delve into the complexities of collective cognition, LLM training, knowledge management, and social interaction, they are interacting with content that has itself been influenced by the advanced technologies under discussion.

This recursive element of the course design illustrates the dynamic and evolving interaction between human intellect and artificial intelligence. It's an embodiment of the

dualities and partnerships that can emerge when the creative capacities of humans are augmented by the meticulousness of machine intelligence. This partnership is indicative of the immense possibilities and responsibilities that come with the integration of such technologies into the fabric of our digital era. Understanding, leveraging, and steering these advancements remain a central theme and imperative throughout this WikiBook.

1.1 Independent Learning

Independent consumption by definition means you can do whatever you want with your time and this book. So do that! Read as much or as little as you like, skip around, and don't be shy about asking questions.

If you are serious about learning the content, you will have to devote significant time. I recommend setting aside consistent time every week, and work through the book section by section. Do the prototypes yourself, and contribute to the wiki book. And email me (am2873@cornell.edu)!

1.2 Teaching the Course

I am developing this textbook in approximately the same structure I imagine one would teach a 9-week intensive (perhaps summer-) course.

Weeks 1-3: Sections 1-3, which are largely theoretical, could be presented in the first three weeks. This makes it a whirlwind tour, but the textbook allows students to dig deeper at their discretion. At the end of the second section *What is Ethical?*, I imagine students would draft an IRB proposal for a intervention they would like to conduct. This serves to focus students on what they'd like to do with the technology before the third theoretical week *How Do We Do It?* and the following technical weeks which prepare the student for their own prototyping.

Weeks 4-5: The next two weeks can be spent on the technical details of LLMs and various other relevant technologies. Students can choose a topic from the textbook to explain to the class, or choose to research a new one and write a wiki chapter.

Weeks 6-8: The next three weeks would involve hands-on prototyping based on the subject matter. This gives a nice fail-fast mentality, and avoids "scope creep," which can easily result in never getting off the ground. This would be wildly benefited by a user-friendly package which includes high-level access to capabilities for everything mentioned in this book.

Week 9: The final week can be used to reflect on the course material, and allow students to present what they were able to do, what challenges they faced, and their ideas for further development and use of these technologies. If they feel they can contribute to the code-base, this would be a good time to submit their pull-requests.

3

---

3   https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates

2 What's Possible?

2.1 Conversation in Theory

I'll keep an older draft[1], just in case.

What do we mean by "meaning"? The subject of this chapter, it's the sort of self-referential question which one may think at first glance to be fundamentally unanswerable. But this apparent circularity has not stopped philosophers and scientists from working through the answer, or stopped humans globally from saying things, and really "meaning" them. Questions of the meaning of "understanding," "explanation," "belief," and "reasoning" have a similar circularity, and answers bearing on one of these concepts help to answer what the others mean. In this chapter we conduct a broad survey of answers to these questions through history, in the hope of assembling a robust answer to this question. Or at least one which will help us understand how to design a system which can understand "for all practical purposes." Without further ado, let's start at the beginning.

2.1.1 Overview of the Fields

Not everything relevant will be treated in this chapter, so we'll begin by giving an alphabetical index of fields possibly relevant to the study of meaning, of the structure of language, meaning, and conversation.

Philosophy of Meaning

- **Cognitive Linguistics:** This area looks at how language and thought interact. It examines how linguistic structures reflect the ways people categorize, conceptualize, and understand the world around them.
- **Existentialism:** Existentialist philosophers like Sartre and Heidegger delve into how meaning is not inherent in the world but is something that individuals must create for themselves, often in the face of an absurd or fundamentally meaningless universe. The tradition of existentialism leads naturally to studies of common understanding, in the taken-for-granted worlds discussed in phenomenology (Heidegger), especially by Schutz.
- **Hermeneutics:** Originally a method of interpreting religious texts, hermeneutics has expanded as a broader methodology for interpreting texts and symbols in general. It's about understanding the meaning behind written and spoken language, often in a historical or cultural context.

---

[1]   https://en.wikibooks.org/wiki/Chatbots%20For%20Social%20Change%2FTheory%20of%20Conversation%2FTheorizing%20Conversation%2Fdraft%201

- **Philosophy of Language:** This branch of philosophy deals extensively with how language and its structure contribute to meaning. It covers topics from the meaning of words and sentences to the nature of communication and understanding.
- **Pragmatics:** As a part of linguistics and semiotics, pragmatics studies how context influences the interpretation of meaning, considering factors like speaker intent, cultural norms, and situational context.
- **Philosophy of Mind:** This branch explores the nature of mental states, consciousness, and how these relate to issues of meaning and understanding. It questions how mental representations carry meaning and how this is communicated.
- **Pragmatism:** Pragmatism, a school of thought in American philosophy, focuses on the practical implications of ideas and beliefs, which is key in understanding how meaning is construed and applied in real-world scenarios.

Sociology of Meaning

- **Critical Theory:** Philosophers like Adorno, Horkheimer, and Habermas in the Frankfurt School have explored how societal structures, power dynamics, and ideologies shape the construction and interpretation of meaning. We will here also include Feminist Sociology in this category, as it makes similar claims that meaning and knowledge are impinged by social structure.
- **Ethnomethodology:** A sociological perspective that examines how people produce and understand social order in everyday life. It's particularly interested in how people make sense of and find meaning in their social world. Has roots in phenomenology, especially Schutz.
- **Interaction Ritual Theory:** Goffman's work focuses on the significance of social interactions and the performative aspects of social life, exploring how meaning and identity are constructed and expressed in everyday rituals and practices.
- **Phenomenology:** Phenomenology, especially as developed by Husserl, focuses on the structures of experience and consciousness. Schutz extended this to the social realm, exploring how individuals understand and ascribe meaning to their experiences and to each other in everyday life.
- **Post-Structuralism:** This movement, with thinkers like Derrida and Foucault, challenges the stability and universality of meaning, suggesting it is fluid, context-dependent, and a product of discourses of power.
- **Semiotics:** This is perhaps the closest field to what you're describing. Semiotics is the study of signs and symbols and their use or interpretation. It's a field that intersects with linguistics, philosophy, and anthropology, focusing on how meaning is created and understood.
- **Social Constructionism:** This perspective, associated with Peter L. Berger and Thomas Luckmann, argues that many of the most familiar aspects of our social world and our understanding of it are not natural, but rather constructed through ongoing social processes.
- **Sociolinguistics:** This interdisciplinary field examines how language use and social factors are interrelated. It looks at how different social contexts and groups shape language, communication styles, and the meanings conveyed through language.
- **Speech Act Theory:** This theory explores how language is used not just for conveying information but for performing actions. Searle, in particular, contributed to the devel-

opment of social ontology, examining how shared understandings and intentions create a framework for meaningful communication and social reality.

- **Symbolic Interactionism:** Mead's theory emphasizes the role of social interaction in the development of the mind and the self. Meaning, in this view, arises from the process of interaction and communication between individuals.

Intersubjectivity

Frames, Speech Acts, Conversational Analysis

Natural Language Processing

Can Machines Understand?

[2]

2.2 Conversation in Practice

Besides the armchair theorists treated in the last section, it will be instructive for us to turn our attention to the strategies which professional practitioners in conversational techniques have uncovered to better perform their duties. Examples include therapists, mediators of conflict and discord, political speech, salesmen, negotiators, educators, coaches and mentors, and healthcare professionals, to name a few. There is no shortage of areas where individuals have to use conversation to get something done, and in these areas we should find their learned strategies to be a helpful guide to the nature of conversational interaction in general.

This section will treat a wide variety of strategies, typologies, and theories developed in different aims over history to solve practical problems with conversation. It is by no means exhaustive, and being in the format of a WikiBook, I encourage enlightened readers to contribute their own knowledge-bases where appropriate.

2.2.1 Carl Rogers and Client-Centered Therapy

Carl Rogers[3] was a pioneering psychologist who developed client-centered therapy[4] in the 1940s and 1950s, a revolutionary approach at the time that emphasized the humanistic aspects of psychology. Rogers believed in the inherent goodness and potential for growth within every individual, a stark contrast to the deterministic views of human behavior prevalent in his time. His work focused on the importance of the therapeutic relationship as a facilitator for personal development and healing. Rogers' theory was built on the idea that people have a self-actualizing tendency - an innate drive towards growth, development, and fulfillment of their potential. By providing a supportive and understanding environment, therapists could help clients unlock this potential.

His approach introduced several core principles that have influenced not just client-centered therapy but also many other counseling theories, including motivational interviewing:

---

2   https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates
3   https://en.wikipedia.org/wiki/Carl_Rogers
4   https://en.wikipedia.org/wiki/Client-centered_therapy

- **Empathy**: Demonstrating a deep, nonjudgmental understanding of the client's experience and feelings.
- **Unconditional Positive Regard**: Offering acceptance and support to the client regardless of their actions or feelings.
- **Congruence**: Being genuine and authentic in the therapeutic relationship, allowing the therapist's true feelings to be evident without overshadowing the client's experience.
- **Self-Actualization**: The belief that every individual has the innate ability to fulfill their potential and achieve personal growth.

Rogers' emphasis on the client's perspective, autonomy, and the therapeutic relationship's quality laid the groundwork for the development of approaches like motivational interviewing, which similarly prioritize empathy, collaboration, and the elicitation of personal motivation for change.

2.2.2 Motivational Interviewing

Miller, W. R., & Rollnick, S. (2013). Motivational interviewing: Helping people change (3rd ed). Guilford Press.

This book is a fantastic introduction and exposition of a method for clinical psychology developed by clinical psychologists William R. Miller[5] and Stephen Rollnick[6], called Motivational Interviewing[7], or MI. Although the Wikipedia article, and certainly the textbook, offer a comprehensive summary of this strategy, it is worth summarizing some of the main principles here for reference and comparison to other conversational strategies.

- **Collaboration over Confrontation**: Emphasizes a partnership that respects the client's autonomy rather than confrontation. It's about listening, understanding, guiding, and respecting the client's perspective.

- **Drawing Out Rather Than Imposing Ideas**: Focuses on evoking the client's motivations and commitments to change, rather than imposing solutions.

- **Focus on Ambivalence**: Recognizes that ambivalence about change is natural. The method works through this ambivalence by exploring the client's conflicting feelings about change.

- **Evoking Change Talk**: Guides clients towards expressing their own arguments for change, recognizing and eliciting the client's reasons for and benefits of change.

- **Responding to Resistance**: Teaches practitioners to view resistance as a signal for more exploration and to work with it rather than confronting it directly.

- **Four Processes of MI**:
  - Engaging[8]: Establishing a working relationship.
  - Focusing[9]: Identifying the direction for change.

---

5 https://en.wikipedia.org/wiki/William%20R.%20Miller

6 https://en.wikipedia.org/wiki/Stephen%20Rollnick

7 https://en.wikipedia.org/wiki/Motivational%20Interviewing

8 https://en.wikipedia.org/wiki/Motivational_interviewing%23Engagement

9 https://en.wikipedia.org/wiki/Motivational_interviewing%23Focusing

- Evoking[10]: Eliciting the client's motivations for change.
- Planning[11]: Developing a commitment to change and planning for action.

- **Core Skills (OARS)**: Built on fundamental communication skills:
  - **Open-ended questions**: To explore the client's thoughts and feelings.
  - **Affirmation**: Recognizing strengths and efforts.
  - **Reflective listening**: Showing understanding of the client's perspective.
  - **Summarizing**: Reflecting back the essence of what the client has expressed.

These strategies are based on the principle that the true power for change lies within the client, and the method's effectiveness has been demonstrated across various settings, including healthcare, addiction treatment, and counseling.

2.2.3 Cognitive Behavioral Therapy

Cognitive Behavioral Therapy[12] (CBT) stands as a counterpoint to client-centered therapy[13] and Motivational Interviewing[14], offering a more structured and directive approach to therapy. Developed by Aaron T. Beck[15] in the 1960s, CBT is based on the premise that dysfunctional thinking leads to negative emotions and maladaptive behaviors. The goal of CBT is to identify, challenge, and modify these negative thoughts and beliefs to improve emotional regulation and develop personal coping strategies that target solving current problems.

CBT's core principles include:

- **Identification of Negative Thoughts**: Clients learn to recognize and identify distorted thoughts that contribute to negative emotions.
- **Challenging Cognitive Distortions**: Through techniques like cognitive restructuring, clients learn to challenge and reframe negative patterns of thought.
- **Behavioral Experiments**: Clients test these new thought patterns in real-life situations, learning to modify their behavior based on healthier cognitions.
- **Skill Development**: CBT emphasizes the development of coping strategies and problem-solving skills to manage future challenges effectively.

While CBT's focus on cognitive processes and direct intervention contrasts with the non-directive, empathetic approach of Rogers' therapy and the ambivalence-focused nature of Motivational Interviewing, it complements these methods by offering an alternative pathway for clients whose needs may be better served by a more structured approach. The diversity of therapeutic approaches highlights the complexity of human psychology and the necessity of tailoring interventions to meet individual client needs.

2.2.4 Crucial Conversations: Navigating High-Stakes Discussions

---

10 https://en.wikipedia.org/wiki/Motivational_interviewing%23Evoking
11 https://en.wikipedia.org/wiki/Motivational_interviewing%23Planning
12 https://en.wikipedia.org/wiki/Cognitive_behavioral_therapy
13 https://en.wikipedia.org/wiki/Client-centered_therapy
14 https://en.wikipedia.org/wiki/Motivational_interviewing
15 https://en.wikipedia.org/wiki/Aaron_T._Beck

Crucial Conversations: Tools for Talking When Stakes are High[16] (2nd Edition), edited by Kerry Patterson, provides a framework for handling conversations where opinions differ, stakes are high, and emotions run strong. The book outlines strategies to ensure these crucial conversations can lead to positive outcomes, rather than misunderstanding and conflict. Here are the key strategies outlined in the book:

- **Start with Heart**: Keeping focused on what you truly want to achieve from the conversation, helping to maintain clarity and prevent emotional derailment.

- **Learn to Look**: Becoming aware of when conversations become crucial and require special attention to maintain constructive dialogue.

- **Make It Safe**: Ensuring the conversation environment is safe by maintaining mutual respect and mutual purpose, allowing all parties to feel comfortable sharing their viewpoints.

- **Master My Stories**: Recognizing and managing the internal stories that influence one's emotional responses and behaviors, aiming to respond more thoughtfully.

- **State My Path**: Communicating your own views clearly and respectfully, sharing both facts and your interpretation, while remaining open to others' inputs.

- **Explore Others' Paths**: Showing genuine curiosity about others' perspectives through asking questions, active listening, and acknowledging their emotions and viewpoints.

- **Move to Action**: Deciding how to proceed after the conversation, including making decisions, determining who will carry them out, and agreeing on follow-up actions.

These principles are designed to facilitate open, honest exchanges that can resolve conflicts, build stronger relationships, and lead to

2.2.5 Schemas for Leading Group Discussions

Group facilitation is a crucial skill for guiding discussions, ensuring productive dialogue, and achieving desired outcomes. Below are central concepts related to effectively leading group discussions within the context of group facilitation:

- **Agenda Setting**: Establishing a clear agenda that outlines topics and objectives to keep the group focused and productive.

- **Neutral Facilitation**: Maintaining neutrality by the facilitator, focusing on process management without advocating for specific outcomes.

- **Active Listening**: Promoting active listening among group members to foster understanding and respect for differing opinions.

- **Conflict Resolution**: Constructively addressing conflicts to find common ground or a mutually agreeable way forward.

- **Encouraging Participation**: Ensuring all participants have the opportunity to contribute, especially encouraging quieter members to speak up.

---

16 https://en.wikipedia.org/wiki/Crucial_Conversations

- **Decision-Making Processes**: Employing decision-making strategies such as consensus, majority vote, or other agreed-upon methods to progress the discussion.

- **Time Management**: Keeping the discussion within the predetermined timeframe to ensure topics are covered effectively.

- **Summarization and Clarification**: Summarizing key points, agreements, and action items for clarity and to confirm next steps.

- **Group Dynamics Awareness**: Navigating the varying personalities, relationships, and power dynamics within the group.

- **Feedback and Evaluation**: Seeking feedback on the facilitation process and outcomes to enhance future group interactions.

- **Ground Rules**: Setting clear guidelines for interaction to promote respectful and productive discussion.

- **Visual Facilitation**: Using visual aids such as whiteboards, charts, or digital platforms to help organize thoughts and illustrate ideas.

Recommended References

- Schwarz, R. (2002). The Skilled Facilitator: A Comprehensive Resource for Consultants, Facilitators, Managers, Trainers, and Coaches. Jossey-Bass.

- Kaner, S. (2014). Facilitator's Guide to Participatory Decision-Making. Jossey-Bass.

- Bens, I. (2005). *Facilitating with Ease!: Core Skills for Facilitators, Team Leaders and Members, Managers, Consultants, and Trainers.* Jossey-Bass.

- Hunter, D., Bailey, A., & Taylor, B. (1995). *The Art of Facilitation: How to Create Group Synergy.* Fisher Books.

2.2.6 Political Conversation

Political conversation is a critical domain of communication that encompasses discussions on societal issues, policy-making, governance, and more. It seeks to foster a structured, respectful, and constructive dialogue among diverse viewpoints to better understand societal challenges and develop actionable solutions. Below are key principles and practices essential to the theory of political conversation:

- **Deliberative Democracy**: Emphasizes informed, rational dialogue in decision-making. Practical applications like deliberative forums and citizen juries aim to enhance the quality of political conversation.

- **Inclusive Dialogue**: Encourages participation from a broad spectrum of voices, including marginalized and minority groups, to ensure comprehensive discussions on societal issues.

- **Civility and Respect**: Promotes decency in political discourse, fostering constructive and less polarized conversations.

- **Fact-Based Dialogue**: Anchors political conversations in verifiable facts to support informed discussions and decision-making.

- **Active Listening**: Urges individuals to listen to and attempt to understand opposing viewpoints, fostering empathy and broader perspectives.

- **Common Ground**: Focuses on identifying areas of agreement to build trust and lay a foundation for addressing contentious issues.

- **Conflict Resolution**: Uses structured dialogue and negotiation techniques to resolve conflicts within political discussions.

- **Transparency and Accountability**: Ensures open communication about decisions, policies, and governmental actions to build trust and foster constructive dialogue.

- **Educational Engagement**: Leverages educational campaigns to inform the public on key issues, promoting a more knowledgeable base for political conversation.

- **Mediated Discussions**: Employs facilitated or mediated discussions to manage contentious issues more productively.

Notable Methods and Initiatives To foster healthier political conversations, several methods and initiatives have proven effective:

- **Citizens' Assemblies and Deliberative Polling**: Forums that gather randomly selected citizens to deliberate on political or societal issues, providing diverse perspectives and grassroots solutions.

- **Online Platforms for Civic Engagement**: Digital platforms, like vTaiwan, use technology to facilitate open, constructive discussions on legislative and societal issues, bridging the gap between citizens and policymakers.

ReferencesThese foundational principles, alongside innovative methods and initiatives, offer a framework for engaging in more constructive, inclusive, and fact-based political conversations:

- Fishkin, J. (2011). *When the People Speak: Deliberative Democracy and Public Consultation.* Oxford University Press.

- Gastil, J., & Levine, P. (2005). *The Deliberative Democracy Handbook: Strategies for Effective Civic Engagement in the Twenty-First Century.* Jossey-Bass.

- Nabatchi, T., Gastil, J., Weiksner, G. M., & Leighninger, M. (Eds.). (2012). *Democracy in Motion: Evaluating the Practice and Impact of Deliberative Civic Engagement.* Oxford University Press.

These principles and references highlight the importance of fostering a healthy political conversation as a cornerstone of democracy, encouraging participation, understanding, and collaboration among all sectors of society.

2.3 Collective Action

Collective action lies at the heart of societal transformation. It represents the concerted effort of individuals and groups to achieve shared objectives. In the context of the digital era, *chatbots*, powered by advanced artificial intelligence, have emerged as key players in

facilitating collective action. They have the potential to act as catalysts in global conversations, influencing social change.

2.3.1 Theoretical Foundations

Collective intelligence is crucial for the success of collective action. The research by Riedl et al. (2021) underscores the importance of equitable participation within groups. In a world increasingly mediated by chatbots, these digital interlocutors can democratize conversations, ensuring diverse voices are heard and valued. Moreover, chatbots can augment the collective decision-making process, contributing to the group's overall intelligence and performance.

Persuasion-oriented methods of collective action, such as those used in political campaigns or advocacy efforts, often risk exacerbating social divides. Here, chatbots can offer a solution by providing platforms for rational and balanced discourse, potentially linking to democratically compiled sources like Wikipedia to counteract misinformation.

Tarrow's contentious politics provide a framework for understanding the dynamics of collective action in confrontational settings. Chatbots, designed with the principles of contentious politics in mind, could facilitate the organization and mobilization of collective actors, thus becoming tools for orchestrating social movements and driving political participation.

2.3.2 Case Studies and Applications

Taiwan's Digital Democracy initiatives, such as vTaiwan and Pol.is, exemplify the successful application of digital platforms in enhancing transparency and fostering collective action. These platforms could be further evolved by integrating chatbots, which would serve as mediators and aggregators of public opinion, enhancing the deliberative process.

Participatory Budgeting represents a direct application of collective action in governance. Chatbots could revolutionize this process by engaging with citizens, gathering their preferences, and helping them understand the implications of budgetary decisions, thus reinforcing the *Lefebvre's "right to the city"* principle.

The **polarization** in the United States demonstrates the potential pitfalls of collective action when it leads to societal division. Chatbots could be programmed to recognize signs of polarization, intervening to introduce alternative perspectives and mediate discussions, thereby promoting social cohesion.

2.3.3 Methods and Strategies in Collective Action

The *modular performances and repertoires* of collective action described by Tarrow can be seen as a precursor to the adaptability required of chatbots in social change scenarios. These digital entities must be capable of operating in varied contexts, advocating for diverse causes, and engaging with different target audiences.

2.3.4 Challenges in Collective Action

Olson's Logic of Collective Action identifies the free-rider problem as a significant challenge in collective endeavors. Chatbots could mitigate this issue by providing personalized incentives for participation and by tracking and rewarding contributions to collective goals.

2.3.5 Impact and Implications

The role of chatbots in collective action extends beyond mere facilitation; they embody the potential for substantial social impact. By leveraging large language models and artificial intelligence, chatbots could unify disparate knowledge bases, serve as platforms for marginalized voices, and contribute to the formation of a more informed and engaged citizenry.

2.3.6 Conclusion

Chatbots are poised to become indispensable in the landscape of collective action. This chapter has delved into the theoretical underpinnings of collective action, explored its application in the digital age, and highlighted the transformative potential of chatbots in promoting social change. As architects of this new digital frontier, we must ensure that the development of chatbots is guided by ethical considerations, inclusivity, and a commitment to fostering constructive dialogue. [17]

2.4 Designing Democracy

Dirk Helbing succinctly *expresses*[18] the need and potentials of a digital upgrade to society. He seems quite interested in the applications of upgraded money (in particular, multi-dimensional money) on the revolution and intelligent control of our incentive systems, and thus our organizational principles. I'm more interested in a deeper question, which he also addresses extensively, especially in his book *Next Civilization* [19]. He discusses such concepts as "computational diplomacy" and "digital democracy".

> To ensure fairness in public sector models, we imagine a future where the people developing these models clearly explain the values behind their assumptions and choices, especially when those choices can affect society. By prioritizing decisions that uphold democratic principles, we can address biases and prevent the unfair outcomes that can result from a top-down approach.
>
> - Helbing (2021) *Next Civilization* [20]

The most comprehensive, recent resource which consolidates the wide variety of prior work in this area, and giving perspective, is

---

17 https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates

18 https://www.youtube.com/watch?v=PJGZpV4PUwY&ab_channel=TEDxTalks

19 Helbing (2021) Next Civilization: Digital Democracy and Socio-Ecological Finance - How to Avoid Dystopia and Upgrade Society by Digital Means ^{https://link.springer.com/book/10.1007/978-3-030-62330-2}

20 Helbing (2021) Next Civilization: Digital Democracy and Socio-Ecological Finance - How to Avoid Dystopia and Upgrade Society by Digital Means ^{https://link.springer.com/book/10.1007/978-3-030-62330-2}

2.5 Free Speech

As the UN Universal Declaration of Human Rights puts it,

> The right to hold own opinions has at least three pillars:
>
> 1. The possibility to get access to relevant information with a reasonable effort (in particular, to the facts, which should be recognizable as such).
> 2. The chance to form an own opinion without being manipulated in that process.
> 3. Sufficient and appropriate opportunities to voice own opinions without fear of being punished, and without censorship.

I see this project as critically aiding in and being guided by this form of the call to the right of free speech and the freedom to hold opinions.

Secondly, As has been put a thousand ways by academics, and as is easily recognized by most in the world at this point, there are significant manipulations of popular opinion happening on a grand scale. In the United States, this involves destruction of popular trust in science, highly funded and sophisticated propaganda campaigns funded by billionaires or conducted by the military industrial complex which re-write textbooks, legislation, and media reports.

And finally, individuals in the U.S. can (for the most part) speak their minds. However, the extent to which they are heard, listened to, recognized, is not a matter of impartial considerate deliberation, of equal weight. Understanding is resource-intensive, and exposure is subject to the whims of the algorithm. In other words, to the features of the presentation, and a winner-takes-all attention space. This leaves most which is spoken, completely unheard. And what bubbles to the top, which rings in our ears, is only a very selective sample of the universe of understanding. Helbing puts this in the language of complex systems as follows:

> [...] asymmetrical interactions may lead to relative advantages of some people over others, the system may get stuck in a local optimum. The global optimum reached when interactions are symmetrical may be better for everyone.

2.6 References

[21]

---

21   https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates

3 What's Ethical?

4 How Do We Do It?

4.1 Beliefs

- Purpose and scope of the chapter.
- The relevance of beliefs in the context of social change and chatbot technology.
- Overview of key themes: Belief formation, technological influence, cultural variations, and redundant belief systems.

Selected bibliography:

What are beliefs?

- Albarracin, M., & Pitliya, R. J. (2022). The nature of beliefs and believing. Frontiers in Psychology, 13, 981925. `https://doi.org/10.3389/fpsyg.2022.981925`
- Schwitzgebel, E. (2022). The Nature of Belief From a Philosophical Perspective, With Theoretical and Methodological Implications for Psychology and Cognitive Science. Frontiers in Psychology, 13, 947664. `https://doi.org/10.3389/fpsyg.2022.947664`

Motivated reasoning

- Bayne, T., & Fernández, J. (2009). Delusion and self-deception: Affective and motivational influences on belief formation. Psychology press.
- Boudon, R. (1994). The Art of Self-Persuasion. Polity Press.
- Boudon, R. (1999). Local vs general ideologies: A normal ingredient of modern political life. Journal of Political Ideologies, 4(2), 141–161. `https://doi.org/10.1080/13569319908420793`
- Ellis, J. (2022). Motivated reasoning and the ethics of belief. Philosophy Compass, 17(6), e12828. `https://doi.org/10.1111/phc3.12828`
- Epley, N., & Gilovich, T. (2016). The Mechanics of Motivated Reasoning. Journal of Economic Perspectives, 30(3), 133–140. `https://doi.org/10.1257/jep.30.3.133`
- Kunda, Z. (1990). The case for motivated reasoning. Psychological Bulletin, 108(3), 480–498. `https://doi.org/10.1037/0033-2909.108.3.480`
- Nisbett, R. E., Ross, L., & Nisbett, R. (1980). Human inference: Strategies and shortcomings of social judgment (Nachdr.). Prentice-Hall.
- Ofsowitz, M. (n.d.). The Psychology of Superstition.

Where do beliefs come from?

- Mannheim, K. (1936). Ideology and Utopia (1979). Routledge.

[1]

---

[1]   `https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates`

4.1.1 Defining Beliefs

Chatbots For Social Change/Beliefs/Defining Beliefs[2]

4.1.2 Redundant Belief Systems

Scientific Belief

The exploration of scientific belief is an essential part of understanding how knowledge evolves and corrects itself over time. Science, by its nature, is a self-correcting enterprise that relies on the concept of falsifiability, as articulated by Karl Popper. Theories within science must be testable and, crucially, capable of being proven wrong. This mechanism ensures that erroneous beliefs are eventually weeded out as empirical evidence mounts against them. However, science is not a straightforward path towards truth. As Thomas Kuhn's analysis of paradigm shifts reveals, science often experiences revolutionary changes, not through a gradual accumulation of knowledge, but through significant leaps that redefine entire fields. For instance, the transition from the Ptolemaic to the Copernican model was not just a simple update but a radical shift in understanding our place in the cosmos.

Resistance to these paradigm shifts is common due to the inertia of established frameworks, vested interests, and the authority of established figures within the scientific community. This resistance indicates that the process of scientific advancement is also a human endeavor, subject to the same social and psychological forces that influence all other areas of human activity. The maxim "Science progresses one funeral at a time," often attributed to Max Planck, underscores the idea that generational changes in the scientific community can be pivotal for the acceptance of new theories and concepts.

Moral Theory and Belief

Moral theory and belief are central to how individuals and societies determine what is right and wrong. The debate between moral realism and anti-realism is foundational to ethics. Moral realists argue that moral truths exist independently of human beliefs and constructions, whereas anti-realists hold that moral truths are human constructs. This debate informs how AI may approach facilitating discussions on moral and ethical issues, as it must navigate the complex terrain of absolute truths versus subjective interpretations.

Further complicating this landscape is the concept of moral relativism, which posits that moral judgments are true or false only relative to specific standpoints, and no single viewpoint is universally privileged over others. This view challenges the AI to remain neutral and respectful of diverse moral perspectives. Additionally, the theory of evolutionary ethics suggests that our moral beliefs may be derived from evolutionary processes that favored cooperative behavior. This perspective implies that moral instincts are not solely derived from rational deliberation but also from inherited social behaviors. Cultural influences also play a significant role in shaping moral beliefs, as ethical frameworks are often deeply entwined with cultural norms and values. An AI mediator operating in this space must therefore be adept at understanding and balancing these varied and often conflicting ethical systems.

---

[2] https://en.wikibooks.org/wiki/%3AChatbots%20For%20Social%20Change%2FBeliefs%2FDefining%20Beliefs

AI Mediation in the Context of Beliefs

AI mediation introduces a novel layer to the discourse on beliefs by providing a platform for objective facilitation of conversations. An AI mediator, devoid of the biases and emotional investments that might affect human mediators, has the potential to bridge communication gaps between individuals. This can lead to the exposure of diverse views, mitigating the effects of echo chambers and confirmation bias. By introducing individuals to a broad spectrum of beliefs, AI can facilitate a more nuanced and comprehensive dialogue.

However, the ethical considerations of AI mediation are paramount. The AI must operate with transparency, ensuring that all parties understand how decisions are made within the system. It must represent diverse viewpoints fairly and avoid any manipulation or undue influence on the participants. The AI should not steer conversations or impose certain views, as this would be ethically problematic. Moreover, the AI's role includes providing continuous feedback and adaptation based on real-world outcomes and user feedback. In this way, AI systems can engage ethically and effectively with the vast space of human beliefs, respecting the complexities of scientific and moral theories while promoting informed and constructive dialogue.

Social Representation Theory

Introduction Social Representation Theory (SRT), developed by Serge Moscovici, centers on the concept that beliefs are pivotal elements of social representations. These representations are a complex web of values, ideas, and practices that enable individuals to make sense of their social realities, navigate their environments, and communicate effectively within their social groups.

Social Representations in Moscovici's View

- Beliefs as Central Elements: In the framework of SRT, beliefs are not isolated thoughts of an individual but are the collective property of a group. They are the building blocks of social representations, which are the shared understandings that inform a community's perception of reality.
- Shared Understanding: The formation of social representations allows communities to form a cohesive interpretation of complex phenomena, simplifying and structuring the social world by providing common ground for interpretation and interaction.
- Function of Social Representations: These shared beliefs and practices establish an order that helps individuals orient themselves and facilitate communication within the community, offering a shared code for classifying and naming the world around them.

Not All Beliefs are Social Representations

- Individual vs. Collective: There's a clear distinction between personal beliefs and the collective beliefs that constitute social representations. Individual beliefs may or may not align with the broader social narratives of a community.
- Evolving Representations: Social representations are dynamic constructs that evolve as societies change and as new beliefs become integrated into the collective understanding of a group.

- Dynamic Nature: The evolving nature of social representations ensures that they remain relevant and reflective of the society's current state, accommodating new beliefs and information as society progresses.

Some Basic Findings

- **Dual Process of Social Representations:**
  - Finding/Conclusion: Moscovici introduced the dual processes of anchoring and objectification as the mechanisms by which social representations are formed.
  - Explanation: Anchoring is the process of assimilating new information by placing it within familiar contexts, whereas objectification transforms abstract concepts into something concrete, making them more understandable and relatable.
- **Core and Peripheral System:**
  - Finding/Conclusion: The core elements of social representations are stable and central to group identity, while the peripheral elements are more susceptible to change and adaptation.
  - Explanation: This structural differentiation within social representations explains how they can maintain stability over time while also adapting to new circumstances and information.
- **Social Representations and Social Identity:**
  - Finding/Conclusion: Social representations are instrumental in forming and sustaining a group's social identity.
  - Explanation: They provide the shared beliefs and values that delineate in-group and out-group boundaries, fostering a sense of belonging and collective identity among group members.
- **Influence of Communication:**
  - Finding/Conclusion: The dissemination and shaping of social representations are heavily influenced by the means of communication, especially the mass media.
  - Explanation: The framing of issues and the portrayal of events in the media can substantially shape public understanding and belief about those issues.
- **Resilience of Stereotypes:**
  - Finding/Conclusion: The persistence and resistance of stereotypes to change can be understood through the lens of social representations.
  - Explanation: Stereotypes are a form of social knowledge deeply anchored in the collective beliefs of a group, making them resistant to contradictory evidence.
- **Role in Social Change:**
  - Finding/Conclusion: Social representations serve not only as a reflection of societal values and norms but also as catalysts for social change.
  - Explanation: As social conditions evolve, new representations can challenge existing beliefs, leading to transformations in societal practices and norms.

SRT's application across diverse domains from health to intergroup relations highlights its significance in understanding the interplay between individual beliefs and collective social understandings.

Abric's Evolutionary Theory

Jean-Claude Abric's evolutionary theory provides a nuanced understanding of how social representations are maintained within societies, distinguishing between core and peripheral elements that comprise and stabilize these representations.

Core and Peripheral System

- Core System:
  - Stability and Resistance to Change: The core is composed of fundamental beliefs tied to the group's collective memory and history, offering stability and resistance to change.
  - Historical and Cultural Anchoring: Core elements are deeply rooted in the group's identity and are essential in giving meaning to the representation.
- Peripheral System:
  - Flexibility and Adaptability: Peripheral elements allow the representation to adapt to new information or contexts without challenging the core's integrity.
  - Heterogeneity and Individual Variance: These elements can differ within the group, providing space for individual nuances and interpretations.
- Functions and Interplay:
  - Continuity and Relevance: The core ensures continuity, while the peripheral adapts to maintain the representation's relevance in changing environments.
  - Negotiation of Meaning: The dynamic between the core and peripheral allows for a balance between shared understanding and individual adaptability.

Application and Example Using the social representation of "marriage" within a traditional culture to exemplify Abric's theory:

- Core System:
  - Unchanging Definition: The core definition, such as marriage being a sacred union between a man and a woman, remains largely constant over time.
  - Historical and Cultural Significance: These beliefs are historically and culturally significant, providing social stability and identity.
- Peripheral System:
  - Evolution of Practices: Peripheral elements, like wedding rituals or acceptance of modern practices, showcase adaptability while respecting the core definition.
  - Variability and Adaptation: The peripheral system's flexibility allows the representation to incorporate new societal trends and values.
- Interplay between Core and Peripheral Systems:
  - Speciation of Belief Systems: The interaction can lead to the evolution of subgroups with distinct beliefs, akin to speciation.
  - Societal Implications: Shifts in peripheral elements, influenced by societal changes, can challenge core elements, potentially leading to significant shifts in social representations.

Dynamics of Core and Peripheral Elements The movement between core and peripheral elements reflects the fluid nature of social representations and their susceptibility to change due to societal shifts or internal group dynamics.

- Movement Between Core and Periphery:
  - Peripheral to Core: Beliefs once considered peripheral can gain prominence and integrate into the core.
  - Core to Peripheral: Core beliefs can become peripheral as societal values evolve.

- Speciation of Belief Systems:
  - Emergence of Subgroups: Differences in core beliefs can lead to the formation of distinct subgroups within a larger tradition.
  - Challenges of Reconciliation: Reconciling divergent belief systems, especially when core beliefs are at odds, can be difficult and lead to conflict.

Core Beliefs and Action The relationship between core beliefs and action, as well as their visibility within the group, plays a critical role in defining group membership and enforcing norms.

- Core Beliefs and Group Identity:
  - Core beliefs often dictate group behaviors and serve as identity markers, with deviations potentially leading to sanctions.
- Peripheral Beliefs and Individuality:
  - Peripheral beliefs, being more individualized, might influence personal behavior but are less likely to result in group-level enforcement.
- Reaction to Belief Violations:
  - Violations of core beliefs elicit strong reactions as they threaten group cohesion, whereas peripheral beliefs allow for more tolerance and adaptability.

Inclusivity and Social Representations The pursuit of an inclusive society may lead to the redefinition of core beliefs, promoting values of diversity and openness.

- Redefining Core Beliefs:
  - Traditional core beliefs may evolve to prioritize inclusivity, potentially challenging the existing group's norms and leading to resistance.
- Inclusivity as the New Core:
  - The principle of inclusivity could become the foundational core of an open society, with non-inclusive beliefs becoming peripheral.
- Societal Cohesion and the Core:
  - While inclusivity is crucial, a society requires core principles for cohesion; a balance must be struck to ensure societal unity without sacrificing foundational values.

In summary, Abric's theory elucidates how social representations are preserved through a balance between the stable core and the adaptable periphery, reflecting the dynamic nature of societal beliefs and practices.

3

4.2 Reasoning

4.2.1 Introduction

The discourse on artificial intelligence (AI)[4] and its application in society increasingly contemplates the role AI can play in enhancing the quality of public dialogue. A critical

---

3   https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates
4   https://en.wikipedia.org/wiki/Artificial_intelligence

component of this contemplation is the analysis of the logical structure of beliefs and the practical structure of argument. The logical structure pertains to the computational proof of beliefs—formal logic that can be validated through systematic processes. On the other hand, the practical structure of argument deals with the formulation of convincing arguments, the kind that resonate on a practical level with individuals and communities. These twin pillars of logic serve as the foundation for AI systems designed to foster rational, informed, and constructive exchanges. This chapter delves into the intertwining roles of formal and practical logic within AI-mediated discourse systems, examining their functions, benefits, and challenges.

The integration of these structures in AI discourse systems is pivotal for ensuring that conversations are not only based on sound reasoning but also resonate with the communicative norms and psychological realities of human interaction. While the logical structure enforces a level of rigidity and coherence in belief systems, the practical structure facilitates the nuanced and often complex human dimension of persuasion and argumentation. The balance between these aspects is essential for the development of AI systems that can genuinely contribute to social change by enriching public discourse.

4.2.2 Logical Structure of Beliefs

The logical structure of beliefs in AI discourse is integral to providing a clear and consistent framework for discussions. This structure can be broken down into several core components:

Formal Representation

In the realm of AI, beliefs can be encoded using a variety of formal logical systems. This formalization allows for an objective and consistent framework that facilitates the process of reasoning and the evaluation of belief statements for their truthfulness. Computational representations, such as predicate logic or Bayesian networks, serve as the backbone for this objective assessment, ensuring that AI systems can navigate complex belief patterns with precision.

Formal representations also provide the means for AI systems to process and understand the various layers of human belief systems, from the simplest of propositions to the most intricate of hypotheses. This allows AI to engage in discussions that require a deep understanding of the logical dependencies and the hierarchies present within belief systems, enabling them to contribute meaningfully to the discourse.

Consistency and Coherence

One of the fundamental roles of formal logic in AI systems is to ensure the consistency and coherence of belief structures. By employing formal logic, an AI can detect contradictions within a set of beliefs, identifying where a person's belief system may be internally inconsistent. This is critical in maintaining rational discussions where all participants can agree on the logical foundations of the discourse.

Moreover, consistency and coherence are not solely about detecting contradictions; they also involve the ability to deduce and infer new beliefs logically. AI systems, through formal

logic, can thereby extend discussions to new but logically consistent territories, enriching the conversation with new insights and perspectives.

Inference and Deduction

The process of inference and deduction in formal logic allows for the derivation of new knowledge from established beliefs. AI systems, equipped with this logical structuring, can reason out implications that may not be immediately apparent to human interlocutors. This capability is particularly valuable in structured debates or analytical discussions where the progression of ideas is critical.

However, it's important to recognize the limitations inherent in formal logic when applied to AI discourse systems. Human beliefs are often not just logical but also emotionally charged, and context-dependent, which means they can sometimes defy the neat encapsulation of formal logic. Moreover, the implications of Gödel's incompleteness theorems suggest that there are always going to be limitations to what can be proven within any formal system. This underscores the necessity for AI systems to also understand and appreciate the nuances of human belief structures that go beyond the scope of formal logic.

4.2.3 Practical Structure of Argument

Beyond the rigidity of formal logic lies the practical structure of argument, which is central to the effectiveness of persuasion and argumentation in human discourse. This structure is influenced by several factors:

Rhetoric and Persuasion

The art of rhetoric plays a significant role in the practical structure of argumentation. It is not enough for an argument to be logically sound; it must also be persuasive. Aristotle's modes of persuasion—ethos (credibility), pathos (emotional appeal), and logos (logical argument)—are as relevant today as they were in ancient Greece. An AI system that understands and applies these modes can engage more effectively with humans, presenting arguments that are not only logically valid but also emotionally resonant and ethically sound.

The influence of rhetoric on argumentation is profound. It shapes not only how arguments are perceived but also how they are received and accepted by an audience. Persuasive communication requires an understanding of the audience's values, beliefs, and emotional states. AI systems, therefore, must be adept at not only constructing logical arguments but also at delivering them in a manner that is contextually appropriate and compelling.

Cognitive Biases

The understanding and navigation of cognitive biases are also crucial to the practical structure of argument. Humans are often influenced more by factors such as the availability heuristic, where a vivid anecdote can outweigh statistical evidence in persuasive power. An AI system that is sensitive to these biases can better tailor its arguments, anticipating

and addressing the psychological factors that influence human decision-making and belief formation.

Socratic questioning and the framing of information are additional tools in the practical argument structure. By asking probing questions, an AI can lead individuals to reflect and reach conclusions independently, which often leads to more profound insight and belief change. Furthermore, how information is framed—the context and presentation—can significantly influence its reception and interpretation. Recognizing and utilizing framing effects is essential for AI systems designed to engage in meaningful dialogue.

4.2.4 Integrating Logic and Rhetoric in AI Mediation

The integration of formal logic and practical argumentation in AI systems is a delicate balancing act:

Balanced Approach

An AI mediator must find a harmonious balance between rigorous logical evaluation and the nuanced understanding of human persuasion. This involves not only pointing out logical inconsistencies but also appreciating that humans are not solely driven by logic. Convincing individuals often requires more than just logical reasoning—it requires engaging with them on a level that resonates with their personal experiences and emotions.

Ethical considerations are paramount in this balance. AI mediators must always prioritize informed consent, transparency, and the autonomy of the individuals they engage with. There is a fine line between facilitating constructive dialogue and manipulating beliefs. The AI's role should be to assist in the navigation of discussions, providing clarity and insight while respecting each individual's right to hold and express their beliefs.

Educational Role

Moreover, AI systems can take on an educational role, helping individuals to understand logical fallacies, cognitive biases, and the elements of effective argumentation. This educational aspect is not just about imparting knowledge but also about fostering the skills necessary for critical thinking and self-reflection. Through this process, individuals can become more informed and autonomous thinkers, better equipped to engage in productive and rational discourse.

4.2.5 Strategically Identifying and Presenting Contradictions

In settings where participants are prepared for their beliefs to be challenged, such as "devil's advocate" or "debate contest" experiments, the logical structure of beliefs can be used strategically:

Pinpointing Contradictions

An AI system's ability to pinpoint contradictions within an individual's belief system is a powerful tool for stimulating critical thinking and reflection. When participants are open to

having their views examined, these contradictions can serve as catalysts for deeper inquiry and reassessment of one's stance.

Forcing participants to re-evaluate their beliefs through the presentation of logically structured dissonant facts can lead to a more robust defense of their positions or to a productive shift in their perspectives. In debate settings, this dynamic can enhance the quality of the discourse, as participants are encouraged to critically engage with the arguments presented and develop a more refined understanding of the issues at hand.

Setting Clear Boundaries

The establishment of clear boundaries for discourse is another benefit of a strong logical structure. If participants can agree on certain axioms or foundational truths, the debate can focus on the implications and conclusions that logically follow from these premises. This helps to prevent discussions from becoming mired in misunderstandings or irrelevant tangents and instead promotes a focused and productive exchange of ideas.

Highlighting inferential gaps is also crucial. Often, individuals hold beliefs based on incomplete reasoning or insufficient evidence. By logically structuring the argument, an AI system can illuminate these gaps, prompting individuals to seek additional information or to critically evaluate the validity of their conclusions.

Promoting Intellectual Honesty

In environments that encourage the challenge of preconceived notions, the logical structuring of arguments promotes intellectual honesty. Participants are more likely to acknowledge points that are logically indefensible and to respect the strength of well-founded arguments. This intellectual honesty is critical for the integrity of the discourse and for the personal growth of the participants involved.

The educational potential of such engagements is immense. Participants not only learn to appreciate the value of logical reasoning but also become more adept at identifying fallacious arguments and understanding the complex nature of their own and others' beliefs.

Guarding Against Misuse

Despite the potential benefits, there is always a risk that the strategic presentation of dissonant facts could be misused. It is imperative to ensure that the process remains respectful, fair, and aimed at mutual understanding and growth, rather than being used as a means to "win" an argument at the expense of others. The ethical use of logic in discourse is essential for ensuring that the pursuit of truth and understanding is not compromised.

In summary, the integration of the formal and practical aspects of logic into AI-mediated discourse is key to promoting informed, rational, and respectful public dialogue. The logical structure provides a solid framework for discussions, ensuring they adhere to principles of reason and coherence. In contrast, the practical structure addresses the complexities of effective communication, persuasion, and the psychology of belief acceptance. An AI mediator capable of adeptly navigating both realms can thus serve as an effective, ethical, and constructive facilitator of conversations, leading to meaningful social change.

4.2.6 LLMs for Implication Mining

The process of implication mining using large language models (LLMs) is an innovative approach that leverages the advanced capabilities of AI to enrich knowledge bases with logical implications derived from user statements. This method is outlined in several steps:

Isolating Belief Structures

The first stage involves the identification and isolation of belief structures from users, which can be accomplished through:

1. **Subset Selection**: Identifying a specific subset of beliefs from the broader belief structure of one or more users, based on random selection, user input, or thematic relevance.
2. **Statement Aggregation**: Compiling the chosen beliefs into a clear and coherent prompt, ensuring that the LLM can process and understand them effectively.

Querying the LLM

Once the belief structures are prepared, they are presented to the LLM in one of two ways:

1. **Direct Implication Query**: The LLM is asked to deduce direct implications from the aggregated statements, essentially to follow a logical thread to a conclusion.
2. **Open-ended Exploration**: The LLM is given the statements and prompted to generate any interesting or novel observations, leading to potentially broader and more diverse insights.

Handling LLM Responses

Responses from the LLM are critically evaluated and refined through:

1. **Filtering & Validation**: Sifting through the LLM's output to identify valid and relevant implications, which may involve manual review or additional LLM processing.
2. **Database Integration**: Incorporating verified implications into the database, which enriches the existing knowledge base and informs future queries and interactions.

Periodic Exploration

To maintain the relevance and growth of the knowledge base, the system includes mechanisms for ongoing exploration:

1. **Scheduled Implication Derivation**: Regularly querying the LLM with different sets of beliefs to uncover new implications and expand the breadth of the database.
2. **User Feedback Loop**: Engaging users in the validation process of the implications associated with their beliefs, promoting accuracy and user interaction.

This structured application of LLMs not only deepens the database's understanding of existing beliefs but also ensures that the knowledge base is dynamic, evolving, and attuned to the complexities of user-provided data.

4.2.7 Proof-Checking

Let's take a brief dive into each of these systems:

1. **Prolog**
   a) *Description:* Prolog (PROgramming in LOGic) is a logic programming language associated with artificial intelligence and computational linguistics. It operates on the principles of formal logic to perform pattern-matching over natural language parse trees and databases.
   b) *Positives:*
      i. Intuitive Semantics: The "facts" and "rules" structure of Prolog is somewhat intuitive for representing beliefs and deriving conclusions.
      ii. Pattern Matching: Prolog excels in pattern matching, which can be valuable for identifying and working with similar beliefs or conclusions.
      iii. Mature Ecosystem: Prolog has been around for a long time and has a variety of libraries and tools.
      iv. Efficient Backtracking: Can explore multiple proof paths and quickly backtrack when a path doesn't lead to a solution.
   c) *Negatives:*
      i. Not Strictly a Proof Assistant: While Prolog can derive conclusions, it doesn't provide formal proofs in the way that dedicated proof assistants do.
      ii. Performance: For very large databases, Prolog might not be the most efficient choice.
   d) *Example:*

```
likes(john, apple).
likes(mary, banana).
likes(john, banana).

likes(john, X) :- likes(mary, X).  % John likes whatever Mary likes

% Query: What does John like?
?- likes(john, What).
% Output: What = apple ; What = banana.
```

2. **Isabelle**
   a) *Description:* Isabelle is a generic proof assistant, which means it allows mathematical formulas to be expressed in a formal language and provides tools for proving those formulas in a logical manner.
   b) *Positives:*
      i. Robust Proof Assistant: Provides rigorous proofs, ensuring the validity of derived conclusions.
      ii. Strong Typing: Helps catch errors early on, making belief representation more accurate.
      iii. Interactive Environment: The semi-interactive nature allows for human-in-the-loop verification and guidance.
      iv. Supports Higher-Order Logic: Can handle complex logical constructs and relations.
   c) *Negatives:*
      i. Steep Learning Curve: Isabelle has a challenging syntax and requires a deep understanding to use effectively.

      ii. Overhead: Might be overkill for simpler belief systems or when you only need to check implications without detailed proofs.

  d) *Example:*

```
theory Example
imports Main
begin

datatype fruit = Apple | Banana

fun likes :: "(string × fruit) set ⇒ string ⇒ fruit ⇒ bool" where
"likes S p f = ((p,f) ∈ S)"

definition exampleSet :: "(string × fruit) set" where
"exampleSet = {(''John'', Apple), (''Mary'', Banana)}"

lemma John_likes_Apple: "likes exampleSet ''John'' Apple"
using exampleSet_def by auto

end
```

3. **Coq**

  a) *Description:* Coq is a formal proof management system. It provides a formal language to write mathematical definitions, executable algorithms, and theorems together with an environment for semi-interactive development of machine-checked proofs.

  b) *Positives:*

      i. Rigorous Proof Mechanism: Like Isabelle, Coq provides very rigorous proofs.

      ii. Extractable Code: Coq allows for the extraction of executable code from definitions, which can be useful if parts of the belief system are algorithmic.

     iii. Strong Community Support: Has a wide range of libraries and an active community.

     iv. Dependent Types: Can express very intricate relationships and properties.

  c) *Negatives:*

      i. Complexity: Like Isabelle, Coq can be difficult to learn and master.

      ii. Performance: Large-scale proof searches can be time-consuming.

     iii. Interactivity: While it's a strength in some contexts, the need for human-guided proof tactics might not be ideal for fully automated reasoning over beliefs.

  d) *Example:*

```
Inductive fruit :=
| Apple
| Banana.

Definition likes (p : string) (f : fruit) : Prop :=
  match p, f with
  | "John", Apple => True
  | "Mary", Banana => True
  | _, _ => False
  end.

Lemma John_likes_Apple : likes "John" Apple.
Proof.
simpl. trivial.
Qed.
```

4. **Z3**

a) *Description:* Z3 is a high-performance theorem prover developed by Microsoft Research. It's used for checking the satisfiability of logical formulas and can be integrated into various applications, including software verification.

b) *Positives:*

i. High Performance: Built for efficiency and can handle large formulas relatively quickly.

ii. SMT Solver: Works with a variety of theories (like arithmetic, arrays, and bit vectors) which could provide versatility in representing beliefs.

iii. APIs for Multiple Languages: Can be integrated easily into various software frameworks.

iv. Decision Procedures: Automatically decides the satisfiability of statements without needing guided tactics.

c) *Negatives:*

i. Not a Traditional Proof Assistant: While Z3 can tell you if a statement is true or false based on given axioms, it doesn't produce detailed proofs in the same manner as Isabelle or Coq.

ii. Expressivity Limitations: Some complex logical constructs might be harder to represent compared to systems like Coq or Isabelle.

d) *Example:*

```python
from z3 import *

# Define the sorts (data types)
Fruit = DeclareSort('Fruit')
Person = DeclareSort('Person')

# Declare the function likes
likes = Function('likes', Person, Fruit, BoolSort())

# Create the solver
s = Solver()

# John and Mary as constants of sort Person
John, Mary = Consts('John Mary', Person)
Apple, Banana = Consts('Apple Banana', Fruit)

# Assert the statements
s.add(likes(John, Apple))
s.add(likes(Mary, Banana))

# Check if John likes Apple
print(s.check(likes(John, Apple)))  # Output: sat (satisfiable)
```

These examples are quite simplistic, and in practice, these tools can handle and are used for much more sophisticated tasks. However, they should provide a basic idea of how each system looks and operates.

4.2.8 Conclusion

The chapter on "Formal and Practical Logic" in the context of "Chatbots for Social Change" underscores the importance of these two intertwined aspects of logic in shaping AI-mediated discourse. It posits that while the logical structure of beliefs lays the groundwork for rational discussions, the practical structure of argumentation brings the human element to the forefront of AI interactions. The challenge for AI systems lies in seamlessly integrating

these structures to promote dialogue that is not only intellectually rigorous but also socially and emotionally engaging. By doing so, AI has the potential to significantly contribute to social change by elevating the quality of public discourse and fostering a more informed, rational, and empathetic society.

5

5   https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates

5 Machine Understanding

5.1 Original chapter organization

5.1.1 Resources

Introductory videos

- *What are LLMs by Google*[1] - A simple 5 minute introduction to what LLMs do.
- *What are LLMs by Apple (WWDC)*[2] - A great introduction to embeddings, tagging, and LLMs, in the first 6m. Then moves to the main topic, multilingual models.
- *Build an LLM from scratch*[3] - A good video introduction, paired with a blog, conceptualizing the process of going from zero to full-scale LLM ($1M later!).
- Microsoft's infrastructure for training chatGPT[4], and similar incredibly high-throughput applications.
    - Uses DeepSpeed[5] to optimize the training
    - InfiniBand[6] for incredible network throughput
    - ONNX[7] to move networks around (also see CRIU[8] for an idea of how checkpointing works])
    - NVIDIA's tensor cores[9] instead of GPUs for compute

Lectures

- *Introduction to Neural Networks*[10] - Stanford CS224N. A nice mathematical overview of Neural Networks.
- *Scaling Language Models*[11] - Stanford CS224N Guest Lecture. A general outlook of the rise in abilities of LLMs.
- *Building Knowledge Representation*[12] - Stanford CS224N Guest Lecture. Very useful to understand our methods for vector retrieval, but from a more general perspective.
    - Dot product has fast nearest-neighbor search algorithms (sub-linear).

---

1   https://www.youtube.com/watch?v=iR2O2GPbB0E
2   https://developer.apple.com/videos/play/wwdc2023/10042/
3   https://www.youtube.com/watch?v=ZLbVdvOoTKM
4   https://www.youtube.com/watch?v=Rk3nTUfRZmo
5   https://github.com/microsoft/DeepSpeed
6   https://en.wikipedia.org/wiki/InfiniBand
7   https://en.wikipedia.org/wiki/Open%20Neural%20Network%20Exchange
8   https://criu.org/Main_Page
9   https://www.nvidia.com/en-us/data-center/tensor-cores/
10 https://www.youtube.com/watch?v=X0Jw4kgaFlg
11 https://www.youtube.com/watch?v=UFem7xa3Q2Q
12 https://www.youtube.com/watch?v=4ynrGLIuPv4

- Re-ranking is often necessary, because the dot-product is not necessarily so expressive.
- Not all sets of vectors are easily indexed, "pathological," and to improve performance it can be beneficial to "spread them out," e.g.
- *Socially Intelligent NLP Systems*[13] - Nice! A deep-dive on how society impinges on language, and how that buggers up our models.
- LangChain vs. Assistants API[14] - a nice overview of two interfaces for deeper chatbot computation.
- Emerging architectures for LLM applications[15], a look from the enterprise side, of architectures and their use. Covers RAG with vector search, Assistants, and general workflow of refining LLM models.
- GPT from scratch[16] - fantastic introduction to chatGPT, understanding exactly how it works (Torch), by Andrej Karpathy.

Courses

- *NYU Deep learning*[17] - A fully-fledged online course, including many advanced topics, including Attention and the Transformer, Graph Convolutional Networks, and Deep Learning for Structured Prediction, for instance.
- The textbook Understanding Deep Learning[18] has accompanying exercises, and lecture slides

Textbooks

- A Hacker's Guide to Language Models[19], by Jeremy Howard
- Practical Deep Learning for Coders[20]
- Understanding Deep Learning[21], with this[22] nice introduction talk
- Dahl, D. A. (2023). *Natural language understanding with Python*: Combine natural language technology, deep learning, and large language models to create human-like language comprehension in computer systems. Packt Publishing.
- Sinan Ozdemir. (2023). *Quick Start Guide to Large Language Models*: Strategies and Best Practices for Using ChatGPT and Other LLMs. Addison-Wesley Professional.
- Zhao, W. X. et al. (2023). A Survey of Large Language Models (arXiv:2303.18223). arXiv[23].
  - Not exactly a textbook, but at 122 pages of dense referenced material it packs a punch, and should not be considered a resource to be consumed in one sitting.

---

13 https://www.youtube.com/watch?v=iXjtJmUQBZk
14 https://www.youtube.com/watch?v=OXruvSd-Tk8
15 https://www.youtube.com/watch?v=Pft04KLw5Lk
16 https://www.youtube.com/watch?v=kCc8FmEb1nY
17 https://atcold.github.io/NYU-DLSP20/en/week01/01/
18 https://udlbook.github.io/udlbook/
19 https://www.youtube.com/watch?v=jkrNMKz9pWU
20 https://course.fast.ai/
21 https://udlbook.github.io/udlbook/
22 https://www.youtube.com/watch?v=sJXn4Cl4oww&ab_channel=MachineLearningStreetTalk
23 http://arxiv.org/abs/2303.18223

5.1.2 Reranking

Reranking in the context of information retrieval is a two-step process used to enhance the relevance of search results. Here's how it typically works:

- First-Stage Retrieval: In the initial phase, a broad set of documents is retrieved using a fast and efficient method. This is often done using embedding-based retrieval, where documents and queries are represented as vectors in a multi-dimensional space. The aim here is to cast a wide net and retrieve a large candidate set of documents quickly and with relatively low computational cost.
- Second-Stage Reranking: The documents retrieved in the first stage are then re-evaluated in the second stage to improve the precision of the search results. This stage involves a more computationally intensive algorithm, often powered by a Language Model (like an LLM), which takes the context of the search query more thoroughly into account. This step reorders (reranks) the results from the first stage, promoting more relevant documents to higher positions and demoting less relevant ones.

The reranking step is a trade-off between the relevance of the search results and the computational resources required. By using it as a second stage, systems aim to balance the speed and efficiency of embedding-based retrieval with the depth and relevance of LLM-powered retrieval. This combined approach can yield a set of results that are both relevant and produced within an acceptable timeframe and cost.

5.1.3 Contradiction Detection

The Stanford Natural Language Processing Group has worked on detecting contradictions in text and has created contradiction datasets for this purpose. They have annotated the PASCAL RTE datasets for contradiction, marked for a 3-way decision in terms of entailment: "YES" (entails), "NO" (contradicts), and "UNKNOWN" (doesn't entail but is not a contradiction). Additionally, they have created a corpus where contradictions arise from negation by adding negative markers to the RTE2 test data and have gathered a collection of contradictions appearing "in the wild".

-
  URAY, MARTIN (2018). "EXPLORING NEURAL MODELS FOR CONTRADICTION DETECTION". UNKNOWN TEMPLATE [[Template:cite journal
  cite journal]]: Cite journal requires |journal= (help[24])Category:CS1 errors: missing periodical[25]

-
  TAWFIK, NOHA S.; SPRUIT, MARCO R. (2018). "AUTOMATED CONTRADICTION DETECTION IN BIOMEDICAL LITERATURE". IN PETRA PERNER (ED.) (ED.). *Machine Learning and Data Mining in Pattern Recognition*[26]. VOL. 10934. CHAM: SPRINGER INTERNATIONAL PUBLISHING. PP. 138–148. ISBN[27] 978-3-319-96135-4 978-3-319-96136-1[28]. RETRIEVED 2023-11-16. UNKNOWN TEMPLATE [[Template:cite book

---

24 https://en.wikibooks.org/wiki/Help%3ACS1%20errors%23missing_periodical
25 https://en.wikibooks.org/wiki/Category%3ACS1%20errors%3A%20missing%20periodical
26 http://link.springer.com/10.1007/978-3-319-96136-1_12
27 https://en.wikibooks.org/wiki/ISBN%20%28identifier%29
28 https://en.wikibooks.org/wiki/Special%3ABookSources%2F978-3-319-96135-4%20978-3-319-96136-1

cite book]]: |editor= has generic name (help[29]); Check |isbn= value: length (help[30])Category:CS1 errors: generic name[31]Category:CS1 errors: ISBN[32]

- HSU, CHENG; LI, CHENG-TE; SAEZ-TRUMPER, DIEGO; HSU, YI-ZHAN (2021-11-16), *WikiContradiction: Detecting Self-Contradiction Articles on Wikipedia*[33], ARXIV, RE-TRIEVED 2023-11-16

- LI, LUYANG; QIN, BING; LIU, TING (2017-05-24). "CONTRADICTION DETECTION WITH CONTRADICTION-SPECIFIC WORD EMBEDDING"[34]. *Algorithms*. **10** (2): 59. doi[35]:10.3390/a10020059[36]. ISSN[37] 1999-4893[38]. Retrieved 2023-11-16.

- AL JALLAD, KHLOUD; GHNEIM, NADA. "ARNLI: ARABIC NATURAL LANGUAGE IN-FERENCE FOR ENTAILMENT AND CONTRADICTION DETECTION". UNKNOWN TEMPLATE [[Template:cite journal

  cite journal]]: Cite journal requires |journal= (help[39])Category:CS1 errors: missing periodical[40]

[41]

5.2 Introduction to Large Language Models (LLMs)

Let's dive into the world of Large Language Models (LLMs). These are advanced computer programs designed to understand, use, and generate human language. Imagine them as vast libraries filled with an enormous range of books, covering every topic you can think of. Just like a librarian who knows where to find every piece of information in these books, LLMs can navigate through this vast knowledge to provide us with insights, answers, and even generate new content.

How do they achieve this? LLMs are built upon complex algorithms and mathematical models. They learn from vast amounts of text – from novels and news articles to scientific papers and social media posts. This learning process involves recognizing patterns in language: how words and sentences are structured, how ideas are connected, and how different expressions can convey the same meaning.

Each LLM has millions, sometimes billions, of parameters – these are the knobs and dials of the model. Each parameter plays a part in understanding a tiny aspect of language,

---

29 https://en.wikibooks.org/wiki/Help%3ACS1%20errors%23generic_name
30 https://en.wikibooks.org/wiki/Help%3ACS1%20errors%23bad_isbn
31 https://en.wikibooks.org/wiki/Category%3ACS1%20errors%3A%20generic%20name
32 https://en.wikibooks.org/wiki/Category%3ACS1%20errors%3A%20ISBN
33 http://arxiv.org/abs/2111.08543
34 https://www.mdpi.com/1999-4893/10/2/59
35 https://en.wikibooks.org/wiki/doi%20%28identifier%29
36 https://doi.org/10.3390%2Fa10020059
37 https://en.wikibooks.org/wiki/ISSN%20%28identifier%29
38 https://www.worldcat.org/issn/1999-4893
39 https://en.wikibooks.org/wiki/Help%3ACS1%20errors%23missing_periodical
40 https://en.wikibooks.org/wiki/Category%3ACS1%20errors%3A%20missing%20periodical
41 https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates

like the tone of a sentence, the meaning of a word, or the structure of a paragraph. When you interact with an LLM, it uses these parameters to decode your request and generate a response that is accurate and relevant.

One of the most fascinating aspects of LLMs is their versatility. They can write in different styles, from formal reports to casual conversations. They can answer factual questions, create imaginative stories, or even write code. This adaptability makes them incredibly useful across various fields and applications.

LLMs are a breakthrough in the way we interact with machines. They bring a level of understanding and responsiveness that was previously unattainable, making our interactions with computers more natural and intuitive. As they continue to evolve, they're not just transforming how we use technology, but also expanding the boundaries of what it can achieve.

In this chapter, we'll explore the world of Large Language Models (LLMs) in depth. Starting with their basic definitions and concepts, we'll trace their historical development to understand how they've evolved into today's advanced models. We'll delve into the key components that make LLMs function, including neural network architectures, their training processes, and the complexities of language modeling and prediction. Finally, we'll examine the fundamental applications of LLMs, such as natural language understanding and generation, covering areas like conversational agents, sentiment analysis, content creation, and language translation. This chapter aims to provide a clear and comprehensive understanding of LLMs, showcasing their capabilities and the transformative impact they have in various sectors.

5.2.1 Definition and Basic Concepts

Foundations of Neural Networks

To truly grasp the concept of Large Language Models (LLMs), we must first understand neural networks, the core technology behind them. Neural networks are a subset of machine learning inspired by the human brain. They consist of layers of nodes, or 'neurons,' each capable of performing simple calculations. When these neurons are connected and layered, they can process complex data. In the context of LLMs, these networks analyze and process language data. The Structure of Neural Networks in LLMs

- Input Layer: This is where the model receives text data. Each word or character is represented numerically, often as a vector, which is a series of numbers that capture the essence of the word.
- Hidden Layers: These are where the bulk of processing happens. In LLMs, hidden layers are often very complex, allowing the model to identify intricate patterns in language. The more layers (or 'depth') a model has, the more nuanced its understanding of language can be.
- Output Layer: This layer produces the final output, which could be a prediction of the next word in a sentence, the classification of text into categories, or other language tasks.

Training Large Language Models

Training an LLM involves feeding it a vast amount of text data. During this process, the model makes predictions about the text (like guessing the next word in a sentence). It

then compares its predictions against the actual text, adjusting its parameters (the weights and biases of the neurons) to improve accuracy. This process is repeated countless times, enabling the model to learn from its mistakes and improve its language understanding. Parameters: The Building Blocks of LLMs

Parameters in a neural network are the aspects that the model adjusts during training. In LLMs, these parameters are numerous, often in the hundreds of millions or more. They allow the model to capture and remember the nuances of language, from basic grammar to complex stylistic elements. From Data to Language Understanding

Through training, LLMs develop an ability to understand context, grammar, and semantics. This isn't just word recognition, but an understanding of how language is structured and used in different situations. They can detect subtleties like sarcasm, humor, and emotion, which are challenging even for human beings. Generating Language with LLMs

Once trained, LLMs can generate text. They do this by predicting what comes next in a given piece of text. This capability is not just a parroting back of learned data, but an intelligent synthesis of language patterns that the model has internalized.

By understanding these fundamental concepts, we begin to see LLMs not just as tools or programs, but as advanced systems that mimic some of the most complex aspects of human intelligence. This section sets the stage for a deeper exploration into their historical development, key components, and the transformative applications they enable.

5.2.2 Historical Development of LLMs

The journey of Large Language Models began with rule-based systems in the early days of computational linguistics. These early models, dating back to the 1950s and 60s, were based on sets of handcrafted rules for syntax and grammar. The advent of statistical models in the late 1980s and 1990s marked a significant shift. These models used probabilities to predict word sequences, laying the groundwork for modern language modeling.

The 2000s witnessed a transition from statistical models to machine learning-based approaches. This era introduced neural networks in language modeling, but these early networks were relatively simple, often limited to specific tasks like part-of-speech tagging or named entity recognition. The focus was primarily on improving specific aspects of language processing rather than developing comprehensive language understanding.

The introduction of deep learning and word embeddings in the early 2010s revolutionized NLP. Models like Word2Vec provided a way to represent words in vector space, capturing semantic relationships between words. This period also saw the development of more complex neural network architectures, such as Long Short-Term Memory (LSTM) networks, which were better at handling the sequential nature of language.

The introduction of the Transformer model in 2017 was a watershed moment. The Transformer, first introduced in a paper titled "Attention Is All You Need," abandoned recurrent layers in favor of attention mechanisms. This allowed for more parallel processing and significantly improved the efficiency and effectiveness of language models. Rise of Large-Scale Language Models

Following the Transformer's success, there was a rapid escalation in the scale of language models. Notable models include OpenAI's GPT series, Google's BERT, and others like XLNet and T5. These models, with their vast number of parameters (into the billions), demonstrated unprecedented language understanding and generation capabilities. They were trained on diverse and extensive datasets, enabling them to perform a wide range of language tasks with high proficiency. Recent Developments: Increasing Abilities and Scale

The most recent phase in the development of LLMs is marked by further increases in model size and capabilities. Models like GPT-3 and its successors have pushed the boundaries in terms of the number of parameters and the depth of language understanding. These models exhibit remarkable abilities in generating coherent and contextually relevant text, answering complex questions, translating languages, and even creating content that is indistinguishable from human-written text.

5.2.3 Architecture

Large Language Models (LLMs), such as those based on the Transformer architecture, represent a significant advancement in the field of natural language processing. The Transformer model, introduced in the paper "Attention Is All You Need", has become the backbone of most modern LLMs.

The architecture of a Transformer-based LLM is complex, consisting of several layers and components that work together to process and generate language. The key elements of this architecture include:

- **Input Embedding Layer**: This layer converts input text into numerical vectors. Each word or token in the input text is represented as a vector in a high-dimensional space. This process is crucial for the model to process language data.

- **Positional Encoding**: In addition to word embeddings, Transformer models add positional encodings to the input embeddings to capture the order of the words in a sentence. This is important because the model itself does not process words sequentially as in previous architectures like RNNs (Recurrent Neural Networks).
- **Encoder and Decoder Layers**: The Transformer model has an encoder-decoder structure. The encoder processes the input text, and the decoder generates the output text. Each encoder and decoder consists of multiple layers.
  - Each layer in the encoder includes two sub-layers: a multi-head self-attention mechanism and a simple, position-wise fully connected feed-forward network.
  - Each layer in the decoder also has two sub-layers but includes an additional third sub-layer for attention over the encoder's output.
- **Self-Attention Mechanism**: This mechanism allows the model to weigh the importance of different words in the input sentence. It enables the model to capture contextual information from the entire sentence, which is a key feature of the Transformer model.
- **Multi-Head Attention**: This component splits the attention mechanism into multiple heads, allowing the model to simultaneously attend to information from different representation subspaces at different positions.
- **Feed-Forward Neural Networks**: These networks are applied to each position separately and identically. They consist of fully connected layers with activation functions.

- **Normalization and Dropout Layers**: These layers are used in between the other components of the Transformer architecture to stabilize and regularize the training process.
- **Output Layer**: The final decoder output is transformed into a predicted word or token, often using a softmax layer to generate a probability distribution over possible outputs.

The Transformer architecture is highly parallelizable, making it more efficient to train on large datasets compared to older architectures like RNNs or LSTMs. This efficiency is one of the reasons why Transformer-based models can be scaled to have a large number of parameters and process extensive language data.

For more detailed information on the Transformer architecture, see the WikiPedia Transformer page[42].

Neural Network Architectures

Training and Data Requirements

Language Modeling and Prediction

5.2.4 Basic Applications of LLMs

Natural Language Understanding

Conversational Agents

Sentiment Analysis

Natural Language Generation

Content Creation

Language Translation

---

42 https://en.wikipedia.org/wiki/Transformer%20%28machine%20learning%20model%29

5.3 Technical Aspects of LLMs

5.3.1 Architectures and Models

Transformer Models

RNNs, LSTMs, and GRUs

5.3.2 Training Large Language Models

Data Preparation and Cleaning

Supervised and Unsupervised Learning Methods

Challenges in Training LLMs

5.3.3 Evaluating LLMs

Performance Metrics

Benchmarks and Testing

5.4 Advanced Techniques and Methodologies

5.4.1 Chain-of-Thought Prompting

Graph of Thoughts paper[43], which has a nice exposition of chain-of-thought, and tree-of-thought prompting.

5.4.2 Zero-Shot and Few-Shot Learning

5.4.3 Retrieval-Augmented Generation (RAG)

5.4.4 Prompt Engineering and Tuning

5.4.5 Hybrid Models

5.4.6 Self-Improving Models

5.4.7 Explainable AI (XAI) Techniques

5.4.8 Contextual Understanding and Memory

5.4.9 Ethical and Bias Mitigation Techniques

5.5 Advanced Applications of LLMs

Hadi et al.'s (2023) 44-page survey [44] offers a solid and recent resource here.

---

43 https://arxiv.org/pdf/2308.09687.pdf

44 Hadi et al. (2023). *Large Language Models: A Comprehensive Survey of its Applications, Challenges, Limitations, and Future Prospects.* TechRxiv, full download available online ^{https://www.techrxiv.org/users/618307/articles/682263-large-language-models-a-comprehensive-survey-of-its-applications-challenges-limitations-and-future-prospects}

5.5.1 Enhancing Creativity and Innovation

Artistic and Literary Creation

Design and Engineering

5.5.2 Complex Problem Solving and Reasoning

Advanced Analytics and Data Interpretation

Decision Support Systems

5.5.3 Personalized Education and Training

Adaptive Learning Systems

Interactive Educational Content

5.5.4 Healthcare and Medical Applications

Medical Research and Drug Discovery

Patient Interaction and Support Systems

5.5.5 Business and Industry Innovations

Automated Business Intelligence

Personalized Marketing and Customer Insights

5.6 Challenges and Future Directions

5.6.1 Scalability and Environmental Impact

5.6.2 Ethical Considerations and Bias

5.6.3 Achieving Generalization and Robustness

5.6.4 Future Trends in LLM Development

Integration with Other AI Technologies

Enhancing Interactivity and Personalization

5.7 Conclusion

6 Implementation and Execution

6.1 Statement Embedding Models

For generating embeddings of statements that can assess the similarity in meaning between two statements, several state-of-the-art, open-source algorithms and tools are available:

OpenAI's Embedding Models[1]OpenAI offers embedding models that are particularly tuned for functionalities such as text similarity and text search. These models receive text as input and return an embedding vector that can be utilized for a variety of applications, including assessing the similarity between statements.

Spark NLP[3]This open-source library provides a suite of transformer-based models, including BERT and Universal Sentence Encoder, which are capable of creating rich semantic embeddings. The library is fully open-source under the Apache 2.0 license.

To use Spark NLP you need the following requirements:

- Java 8 and 11
- Apache Spark 3.5.x, 3.4.x, 3.3.x, 3.2.x, 3.1.x, 3.0.x

GPU (optional): Spark NLP 5.1.4 is built with ONNX 1.15.1 and TensorFlow 2.7.1 deep learning engines. The minimum following NVIDIA® software are only required for GPU support:

- NVIDIA® GPU drivers version 450.80.02 or higher
- CUDA® Toolkit 11.2
- cuDNN SDK 8.1.0

- There is a massive text embedding benchmark (MTEB [5]) which should help us determine which embedding algorithm to use.

6.2 Vector Similarity Search

6.2.1 LLMRails

The MTEB led me to the ember-v1 model by llmrails, because of its success on the **SprintDuplicateQuestions** dataset. The goal is to embed statements such that statements or questions which are deemed by a community to be duplicates are closest. The dataset compiles marked duplicates from *Stack Exchange*, the *Sprint technical forum website* , and *Quora*.

---

1   Introducing text and code embeddings [2]. OpenAI . Retrieved
3   GPU vs CPU benchmark [4]. Spark NLP . Retrieved
5   MTEB ^{https://huggingface.co/blog/mteb}

LLMrails [6] is a platform that offers robust embedding models to enhance applications' understanding of text significance on a large scale. This includes features like semantic search, categorization, and reranking capabilities.

Pricing: "Elevate your data game with our cutting-edge ChatGPT-style chatbot! All you need to do is link your data sources and watch as our chatbot transforms your data into actionable insights."

LLMRails is revolutionizing search technology, offering developers unparalleled access to advanced neural technology. Providing more precise and pertinent results paves the way for transformative changes in the field of search technology, making it accessible to a wide range of developers.

From the website: "with private invitation, join the LLMRails and start your AI advanture!" How did they get this wrong?

- Embed $0.00005 per 1k tokens
- Rerank $0.001 per search
- Search $0.0005 per search
- Extract $0.3 per document

> **Note:** This service does not give the capabilities I need. It's a bit *too* managed. I just need vector embeddings, and retrieval.

6.2.2 Other Vector Databases as a Service

- **Amazon OpenSearch Service** is a fully managed service that simplifies deploying, scaling, and operating OpenSearch in the AWS Cloud. It supports vector search capabilities and efficient vector query filters, which can improve the responsiveness of applications such as semantic or visual search experiences.
- **Azure Cognitive Search**: This service allows the addition of a vector field to an index and supports vector search. Azure provides tutorials and APIs to convert input into a vector and perform the search, as well as Azure OpenAI embeddings for tasks like document search.

- Zilliz Cloud, powered by the world's most powerful vector database, Milvus, solves the challenge of processing tens of billions of vectors.
  - Zilliz has a 30-day free trial worth $400 of credits. 4CUs
  - Pricing: Zilliz Cloud Usage (Each unit is 0.1 cent of usage) $0.001 / unit
- A more comprehensive list, Awesome Vector Search, on GitHub.[7]
  - For cloud services, they list Zilliz first, then Relevance AI, Pinecone, and MyScale.
- Graft somehow came up
  - It is *extremely* expensive, $500/month for 10,000 data points. Unlimited datapoints at $5k/month...
  - Perhaps it's more managed than a Zilliz, or that's just what the infrastructure costs either way?

---

6   llmrails ^{https://www.llmrails.com/}
7   Awesome Vector Search ^{https://github.com/currentslab/awesome-vector-search}

- High price could also be an indication of the value of this sort of technology (they also do the embedding & document upload for you).

6.2.3 Open Source Models

- Milvus is a "vector database built for scalable similarity search, Open-source, highly scalable, and blazing fast." Seems perfect. They have a managed version, but I'm not sure it's necessary now. [8]
- **Elastic NLP: Text Embeddings and Vector Search**: Provides guidance on deploying text embedding models and explains how vector embeddings work, converting data into numerical representations[9].
- **TensorFlow Recommenders' ScaNN**[10] TensorFlow provides an efficient library for vector similarity search named ScaNN. It allows for the rapid searching of embeddings at inference time and is designed to achieve the best speed-accuracy tradeoff with state-of-the-art vector compression techniques.
- Other notable vector databases and search engines include **Chroma, LanceDB, Marqo, Qdrant, Vespa, Vald, and Weaviate**, as well as databases like **Cassandra, Coveo, and Elasticsearch OpenSearch** that support vector search capabilities.

Milvus Benchmark

Milvus has conducted benchmarks, which should give us an idea of overall cost, and how much we can scale before buckling.

- *CPU*: An Intel(R) Xeon(R) Gold 6226R CPU @ 2.90GHz. This is a high-end server-grade processor suitable for demanding tasks. It belongs to Intel's Xeon scalable processors, which are commonly used in enterprise-level servers for their reliability and performance.
- *Memory*: 16*\32 GB RDIMM, 3200 MT/s. This implies that the server has 16 memory slots, each with a 32 GB RDIMM (Registered DIMM) module, totaling 512 GB of RAM. The memory speed is 3200 MT/s (MegaTransfers per second), which indicates how fast the memory can operate.
- *SSD*: SATA 6 Gbps. This indicates that the server uses a Solid State Drive connected through a SATA interface, with a transfer rate of 6 Gigabits per second. SSDs are much faster than traditional hard drives and are preferred for their speed and reliability.

To find an approximate AWS EC2 equivalent, we would need to match these specs as closely as possible. Given the CPU and memory specifications, you might look into the EC2 instances that offer Intel Xeon Scalable Processors (2nd Gen or 3rd Gen) and the ability to configure large amounts of memory.

A possible match could be an instance from the m5 or r5 families, which are designed for general-purpose (m5) or memory-optimized (r5) workloads. For example, the r5.12xlarge instance provides 48 vCPUs and 384 GiB of memory, which, while not an exact match to your specs (since it has less memory), is within the same performance ballpark.

However, keep in mind that AWS offers a wide range of EC2 instances, and the actual choice would depend on the specific balance of CPU, memory, and I/O performance that

---

8   Milvus homepage ^{https://milvus.io/}
9   Elastic ^{https://www.elastic.co}
10   Efficient serving [11]. TensorFlow Recommenders . Retrieved

you need for your application. Also, pricing can vary significantly based on region, reserved vs. on-demand usage, and additional options like using Elastic Block Store (EBS) optimized instances or adding extra SSD storage.

Using the AWS pricing calculator, this amounts to $3 hourly.

- **Search** - (cluster w/1) 7k to 10k QPS @ 128 dimensions, (standalone w/1) 4k to 7.5k QPS
- **Scalability**
  - From 8-16 CPU cores, it doubles. After that it kinda less-than-doubles
  - Going from 1-8 replicas changes QPS from 7k to 31k, and over doubles available concurrent queries (to 1200)

There are 3600 seconds in an hour, so $PQ = $3 / (7k * 3600) = $0.000000119 per query...

6.3 Large Language Models

A useful article, comparing open-source LLM models, was published here, in Medium[12].

6.3.1 LLMs In Hosted Environments

Model	Cost per 1M input tokens	Cost per 1M output tokens	Additional Notes
AI21Labs Jurassic-2 Ultra	$150	$150	Highest quality
AI21Labs Jurassic-2 Mid	$10	$10	Optimal balance of quality, speed & cost
AI21Labs Jurassic-2 Light	$3	$3	Fastest and most cost-effective
AI21Labs Jurassic-2 Chat	$15	$15	Complex, multi-turn interactions Free for $1000 in usage.
Anthropic Claude Instant	$1.63	$5.51	Low latency, high throughput
Anthropic Claude 2.0, 2.1	$8	$24	Best for tasks requiring complex reasoning
Cohere Command	$1.00	$2.00	Standard offering
Cohere Command Light	$0.30	$0.60	Lighter version
Google Bard	Free (although likely limited)		Requires a Google account
GPT-4 Turbo (gpt-4-1106-preview)	$10	$30	
GPT-4 Turbo (gpt-4-1106-vision-preview)	$10	$30	
GPT-4	$30	$60	
GPT-4-32k	$60	$120	
GPT-3.5 Turbo (gpt-3.5-turbo-1106)	$1.00	$2.00	
GPT-3.5 Turbo (gpt-3.5-turbo-instruct)	$1.50	$2.00	

---

[12] `https://medium.com/illuminations-mirror/zephyr-7b-%CE%B1-vs-llama-2-70b-vs-mistral-7b-unraveling-the-future-of-ai-language-models-a34d95968f40`

6.3.2 LLMs On Your Own Hardware

From the model card: "Llama 2 is an auto-regressive language model that uses an optimized transformer architecture. Llama 2 is intended for commercial and research use in English. It comes in a range of parameter sizes—7 billion, 13 billion, and 70 billion—as well as pre-trained and fine-tuned variations."

It turns out that you have to ask Microsoft nicely to get access to the parameter sets, agreeing to terms of use.

- It's clear from a little research[13] that running and training locally (I have a 2021 Mac M1) is going to cause a lot of headache.
- AWS Sagemaker seems to be a great option for getting up and running with open source models.
  - Has access to dozens of models of varying sizes through their Jumpstart feature
  - In practice, you say "go" and are dropped right away in a JupyterLab instance.

Hardware requirements of Llama (Nov 2023)

Model	Instance Type	Quantization	# of GPUs per replica	Cost
Llama 7B	(ml.)g5.2xlarge	-	1	$1.52 (ml.)
Llama 13B	(ml.)g5.12xlarge	-	4	$7.09 (ml.)
Llama 70B	(ml.)g5.48xlarge	bitsandbytes	8	$20.36 (ml.)
Llama 70B	(ml.)p4d.24xlarge	-	8	$37.69 (ml.)

Benchmarking AWS SageMaker and Llama

Fortunately, Phil Schmid has conducted[14] thorough benchmarks of different deployments of Llama on SageMaker in AWS. His blog posts in 2023 in particular are an incredible reference for getting started with these LLMs.

To give the most economical example, The g5.2xlarge ($1.52 / hour) can handle 5 concurrent requests delivering 120 tokens of output per second. Incredible! That's $3.50 per 1M tokens. ChatGPT, for comparison, offers gpt-3.5-turbo (the cheapest option) at $0.0020 per 1K tokens, or $2.00 per 1M tokens. Comparable, and not surprising that OpenAI is cheaper.

Let's compare the most expensive to the most sophisticated OpenAI model, GPT-4. Llama 70B runs on a $37.69 server (ml.p4d.24xlarge) serving 20 concurrent requests at 321 tokens/second. That's $10.43 per 1M tokens. For comparison, GPT-4 costs $0.06 per 1K tokens, or $60 per 1M.

It should be noted as well that Phil Schmid was able[15] to get decent performance (15 seconds per thousand tokens generated) for a budget deployment in AWS's new inferentia2 hardware (inf2.xlarge), which costs just $0.75 per hour. That's $550 per month, so better not leave it on, but still. Very cool!

---

13 `https://www.reddit.com/r/LocalLLaMA/comments/16igu6g/apple_coreml/`
14 `https://www.philschmid.de/sagemaker-llama-benchmark`
15 `https://www.philschmid.de/inferentia2-llama-7b#4-deploy-a-real-time-inference-endpoint-on-amazon-sagemaker`

He trains[16] a 7B parameter Mistral model using ml.g5.4xlarge ($2.03 / hour). It was able to fine-tune based on 15,001 examples, processed in whole 3 times (epochs), in 3.9 hours, giving an overall cost of <$8.

[17]

16 https://www.philschmid.de/sagemaker-mistral#3-fine-tune-mistral-7b-with-qlora-on-amazon-sagemaker
17 https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates

7 Featured Interviews

7.1 Dirk Helbing

Dirk Helbing[1], an esteemed scholar at the forefront of complex systems, social dynamics, and computational science, has been instrumental in exploring the interplay between technology and society. His recent endeavors in advancing the concept of deliberative democracy, especially through his contributions in "Democracy 2.0," underscore the transformative potential of chatbots in facilitating and enhancing human discourse. His collaboration and insights serve as a cornerstone for this course, which delves into the profound impact of chatbots on our collective future.

7.1.1 Ethical and Societal Implications of Chatbots

Opportunities and Challenges Chatbots offer transformative opportunities for enhancing user interaction with technology, providing a platform for information access and engagement. Yet, these opportunities are accompanied by challenges that are not immediately apparent, necessitating a proactive approach to identify and mitigate potential risks. The development of chatbots should not be driven by the pursuit of an "average" or uniform standard, as this risks the homogenization of society and undermines the robustness that diversity brings to social systems. Instead, chatbots should be designed to reflect a wide array of thoughts, supporting a tapestry of perspectives that enrich the societal discourse.

Ethical Influence and Autonomy The ethical use of chatbots centers around the principles of consent and self-determination. Influencing users without their explicit consent raises ethical concerns, highlighting the need for transparency in how chatbots are used as tools for individual empowerment rather than for manipulation. Preserving user autonomy in decision-making is crucial, ensuring that chatbots support rather than undermine the individual's ability to make informed choices. This is particularly relevant in the context of virtual reality experiments, where the immersive nature of the technology holds significant potential for influencing user perceptions and behaviors.

7.1.2 Chatbots in Conflict Resolution and Social Cohesion

Bridging Perspectives Chatbots have the potential to serve as mediators in conflict resolution by exposing individuals to a multitude of perspectives, thereby fostering a deeper understanding of the issues at hand. This process, while intricate and time-consuming, can be instrumental in identifying the roots of conflicts and facilitating pathways towards resolution. By promoting empathy and understanding, chatbots can contribute to diffusing societal tensions and building a more cohesive social fabric.

---

1   `https://en.wikipedia.org/wiki/Dirk%20Helbing`

Diversity and CustomizationA diverse ecosystem of chatbots is essential to cater to the varied needs and preferences of users. Encouraging the development of multiple chatbot models ensures that no single approach becomes dominant, allowing for a more personalized and user-centric experience. This diversity not only enhances the user experience but also serves as a safeguard against the risks associated with a one-size-fits-all solution, which can lead to large-scale, institutionalized errors.

7.1.3 Security and Long-Term Impact of Chatbots

System Vulnerability and Security The susceptibility of chatbot systems to hijacking and misuse is an ongoing concern that requires constant vigilance and innovation in security measures. As with any system, there is an arms race between those working to maintain stability and those with malicious intent. Drawing from the stabilizing features of scientific inquiry, it is imperative to understand and incorporate similar features into chatbot systems to ensure their resilience and robustness.

Delayed Consequences and Ethical Research The conversation around chatbots also touches on the potential for delayed societal consequences, which may not become evident until years into the future. This underscores the importance of a cautious approach to the deployment of chatbots, with a focus on the long-term implications of their integration into daily life. Ethical research and development in the field of AI must prioritize the advancement of knowledge while being cognizant of the potential for misuse, especially by entities that may disregard ethical considerations for profit or political gain.

7.1.4 Deeper Reflections on the Role of Chatbots

This section was written by GPT-4, and offers some food for thought, extending the conversation with Dirk Helbing.

Balancing Scalability with Individual Needs The scalability of chatbots presents a paradox; while it allows for widespread access and utility, it also risks the loss of individualized responses that cater to specific user contexts. The challenge lies in designing chatbots that can provide the efficiency benefits of scalability while maintaining the ability to personalize interactions. This balance is critical in ensuring that chatbots remain versatile and effective across a broad spectrum of scenarios without compromising the unique needs of each user.

Transparency, Accountability, and Public Engagement Transparency and accountability in chatbot operations are essential to build trust and ensure that users understand the rationale behind the chatbot's responses. Public education and awareness initiatives are equally important, equipping users with the knowledge to critically engage with chatbot technology and understand its limitations. Feedback mechanisms play a pivotal role in this ecosystem, providing a direct channel for users to influence the evolution of chatbot systems and ensuring that developers can respond to concerns and improve the technology continuously.

Ethical Frameworks and Regulatory Considerations The development and deployment of chatbots must be underpinned by robust ethical frameworks that guide their influence on individual decision-making. As chatbots become more integrated into societal

functions, the question of regulation becomes increasingly pertinent. What role should regulation play, and how can it be implemented to foster innovation while protecting individual rights and societal values? These questions highlight the need for a collaborative approach involving policymakers, technologists, and the public to navigate the complex ethical landscape of AI and chatbots.

2

7.2 Jeremy Foote

This article details a conversation between Alec and Jeremy, focusing on their collaborative research into the use of chatbots for mediation and social impact in online communities. It first outlines their individual projects, shared interests, ethical considerations, technical challenges, and future steps in the realm of chatbots and large language models (LLMs).

Alec and Jeremy's dialogue represents a confluence of interdisciplinary research involving communication studies, artificial intelligence, and ethics. Their exchange sheds light on the potential of chatbots to reshape online interactions, the ethical implications of deploying such technologies, and the theoretical frameworks that support their usage.

7.2.1 Jeremy's Reddit Project

Jeremy's initiative aims to engage with Reddit users known for toxic commenting by offering them a chance to converse with a bot. This project is a partnership with a scientist studying toxicity and features three distinct approaches to interaction:

- The default setting without specific behavioral nudges.
- An approach that encourages adherence to Reddit's community norms.
- A narrative storytelling mode prompting users to reflect on their behavior.

The goal is to monitor behavioral changes post-interaction to evaluate the effectiveness of each method.

 User Response Categories During the project, Jeremy observed varied responses from users, which he categorized into:

- Trolls who seek to provoke the bot.
- Justifiers who defend their actions and challenge the bot's contextual understanding.
- Good Faith Interactors who show genuine reflection and sometimes gratitude, raising concerns about the authenticity of their thanks during presentations.

 Project Logistics and Concerns

- The logistical aspects of the project, including the initial focus on text interactions due to the complexities of real-time voice generation.
- The potential necessity of a network of specialized LLMs to handle diverse interaction scenarios.

---

2   https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates

7.2.2 Alec's Perspective and Project

Alec views these interactions as educational, employing varied conversational techniques to instill a lesson. His vision includes:

- Utilizing chatbots to bridge individual perspectives, forming a social semantic database from these conversations.
- Reflecting diverse viewpoints in societal discussions through bot-mediated dialogue.
- Exploring data collection methods, including voice or text input, and the feasibility of debate-style interactions with bots.
- Investigating the technical training of LLMs to support his envisioned semantic database.

Alec's Expanded Vision

- Alec's broader aim involves creating a universal mediator to enhance communication efficiency in society, tackle misinformation, and bridge strategic societal divisions.
- His commitment to a non-interventionist and transparent approach.
- He speculates on scaling the project with philanthropic funding.

Potential Applications and Ambitions

- The concept of an AI politician and organizing collective actions through a chatbot.
- A tool that could mediate conversations, reflecting multiple viewpoints for better societal understanding.

7.2.3 Ethical and Technical Considerations

Both researchers deliberate the importance of transparency when users are conversing with a bot, suspecting that such openness might improve the quality of conversations.

Ethical Justifications

- The ethics of engaging with users, especially vulnerable populations, and the argument for universally respectful engagement.
- The process of obtaining consent, as per the Institutional Review Board (IRB) protocols, via Reddit messages and linking to an informational page.

Technical Aspects

- Discussion of the underlying technology, with Jeremy referencing GPT-3[3], and the importance of aligning the bot to ensure safe and kind responses.
- Alec's emphasis on the technical knowledge required for training LLMs and the potential use of open-source alternatives that could run on personal computing resources.

7.2.4 Social Impact and Theoretical Foundations

The discussion also delves into the social and theoretical implications of using LLMs in online communities, referencing social capital theory and the potential for chatbots to mediate discussions that could lead to greater social cohesion.

Democratic Ideal and Simplified Solutions

---

3   https://en.wikipedia.org/wiki/GPT-3

- The conversation touches on the democratic potential of enabling discussions between individuals with opposing views, and the possibility of chatbots mediating these exchanges to prevent devolution into partisanship.

7.2.5 Further Details

Evaluating Outcomes and Measurements

- Insights into behavior change metrics are shared by Jeremy, with an estimate that around 10% of the users they reach out to agree to interact with the bot, and of these, 70% engage to some degree.

Experiment Design and Reddit's Role

- The design of experiments to engage Reddit users, considering reaching out to Reddit users or moderators for participation without necessarily needing permission from Reddit itself.

Closing the Conversation

The conversation concludes with plans for Alec to submit his IRB proposal and for Jeremy to introduce Alec to a colleague interested in mediation using LLMs. They express a mutual interest in exchanging IRB materials and continuing their discussion on potential collaborations.

Potpourri of Relevant Facts

This section compiles specific facts and figures mentioned during the conversation that provide further depth and clarity to the projects and visions discussed by Alec and Jeremy.

- Jeremy's project interaction rates indicate that approximately 10% of users they reach out to agree to interact with the bot. Of these interactions, about half are considered to be in good faith, with 70% engaging to some degree and a third displaying bad faith from the start.
- Jeremy references the use of GPT-3[4] for his chatbot, but later clarifies that they are using GPT-3.5, which has undergone alignment to ensure safe responses.
- Alec is considering using open-source alternatives to GPT-3 that could run efficiently on a Mac, suggesting a hands-on approach to prototyping and potential deployment.
- Jeremy has contemplated investing in GPU resources to facilitate his project's technological needs.
- Alec and Jeremy discuss the possibility of Alec collaborating with Josh Becker from UCL, a researcher with expertise in mediation, who has also shown interest in using LLMs in this context.
- Alec is preparing to submit an IRB proposal and is considering how to structure the consent process, particularly for vulnerable populations.
- Jeremy explains their approach to obtaining consent for their Reddit study through direct messages that link to an informative page, and they ensure users are aware that participation is limited to those over the age of 18.

---

4   https://en.wikipedia.org/wiki/GPT-3

- Alec's motivation for his project was further reinforced after reading "Dark Money"[5] by Jane Mayer, which details the influence of the Koch brothers on American politics.
- Alec and Jeremy contemplate the potential of LLMs like GPT-3.5 in the facilitation of insightful conversations and their significant contributions to the social sciences.
- They also discuss the logistical challenges involved in real-time voice generation for chatbots and the possibility of defaulting to text messaging if the voice proves too complex.

The facts presented in this section not only contribute to the granularity of the projects discussed but also highlight the nuanced considerations taken by both researchers in their pursuit of innovative applications for chatbots and LLMs in enhancing online community interactions and societal communication.

Links

- Jeremy's Homepage[6]
- Joshua Becker's Homepage[7]

[8]

- My thought process[9] after this talk...

7.3 Belief Change Megastudy (Jan Voelkel et al. 2023)

The paper presents a comprehensive analysis based on a large-scale study (n=32,059) exploring strategies to reduce partisan animosity and anti-democratic attitudes in the United States.

Key Findings

- *Partisan Animosity Reduction*: The study found that interventions emphasizing sympathetic individuals with different political beliefs or highlighting common cross-partisan identities were highly effective in reducing partisan animosity.
- *Impact on Anti-Democratic Attitudes*: Interventions aimed at correcting misperceptions about outpartisans' views significantly decreased support for undemocratic practices and partisan violence.
- *Effective Strategies for Partisan Animosity*: Strategies involving relatable, sympathetic outpartisans and common identities across parties showed the highest effectiveness.
- *Effective Strategies for Undemocratic Practices and Partisan Violence*: Interventions correcting exaggerated stereotypes of other party supporters and highlighting the consequences of democratic collapse were effective. Additionally, showing endorsements of democratic principles by political elites also yielded positive results.

Psychological Dimensions of Political Attitudes

---

5   https://en.wikipedia.org/wiki/Dark_Money_%28book%29

6   https://jeremydfoote.com

7   https://www.mgmt.ucl.ac.uk/people/joshuabecker

8   https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates

9   https://en.wikibooks.org/wiki/%2Fthoughts

- The study reveals that partisan animosity and support for undemocratic practices/partisan violence are distinct dimensions, necessitating tailored approaches for interventions.

Characteristics of Effective Interventions

- Explicitly addressing partisan animosity and undemocratic practices in interventions proved more effective.
- Multifactorial interventions, which employed multiple theoretical mechanisms, were generally more effective than unifactorial ones.
- Higher production quality and engaging content also contributed to the effectiveness of interventions.

Broader Applications

- The findings offer insights for applications on websites and social media platforms to address partisan division and anti-democratic attitudes.
- The study provides a foundation for future research and the development of interventions aimed at strengthening democratic attitudes in the U.S.

These insights provide valuable approaches for addressing the growing partisan divide and threats to democratic principles in America.

7.3.1 Implications

The findings from the "Megastudy identifying effective interventions to strengthen Americans' democratic attitudes" offer profound implications for the logistics, utility, and impact of chatbots in fostering intelligent social action:

- *Empathy and Understanding in Chatbots:* The success of interventions that highlight sympathetic individuals with different political beliefs underscores the importance of empathy and understanding in chatbots. Chatbots, equipped with NLU and LLMs like ChatGPT, can be programmed to represent diverse perspectives and facilitate conversations that foster empathy among users. They can simulate dialogues where users experience differing viewpoints, helping to reduce partisan animosity.

- *Correcting Misperceptions through Information:* The effectiveness of interventions aimed at correcting misperceptions about outpartisans' views highlights a crucial role for chatbots. They can serve as fact-checkers and sources of unbiased information, addressing false beliefs and stereotypes in real-time during conversations. This function is particularly vital in reducing support for undemocratic practices and partisan violence.

- *Tailored Intervention Strategies:* Recognizing the distinct dimensions of partisan animosity and support for undemocratic practices, chatbots can be designed to deliver tailored interventions. They can adapt their responses based on the user's attitudes, providing more personalized and effective engagement.

- *Multifactorial Intervention and Engagement:* The success of multifactorial interventions suggests that chatbots should integrate various strategies – such as providing information, simulating empathetic conversations, and highlighting commonalities – to be more effective. Additionally, engaging content and high production quality are crucial for chatbot interfaces to ensure user retention and effective communication.

- *Broader Social Impact:* Chatbots have the potential to be deployed on a large scale across various platforms, including social media and websites. They can serve as tools for reducing partisan division and strengthening democratic attitudes, reaching a wide audience and facilitating social change.

- *Continuous Learning and Improvement:* Given the rapidly evolving nature of social attitudes and political dynamics, chatbots must be equipped with learning algorithms to continually update their strategies and information. This adaptability ensures they remain effective and relevant.

- *Ethical Considerations and Responsibility:* The deployment of chatbots for social change must be guided by ethical considerations, ensuring they promote healthy democratic values and do not reinforce biases or misinformation. Responsible design and deployment are key to maximizing their positive impact.

In summary, the insights from the megastudy offer valuable guidance for the development of chatbots as tools for social change. By leveraging NLU and LLM capabilities, chatbots can be powerful allies in reducing partisan animosity, correcting misperceptions, and fostering a more empathetic and informed society.

10

---

10 https://en.wikibooks.org/wiki/Category%3ABook%3ASpecial%3AExpandTemplates

8 Contributors

Edits	User
13	Amcgail[1]

1   https://en.wikibooks.org/w/index.php%3ftitle=User:Amcgail&action=edit&redlink=1

List of Figures

- GFDL: Gnu Free Documentation License. `http://www.gnu.org/licenses/fdl.html`

- cc-by-sa-3.0: Creative Commons Attribution ShareAlike 3.0 License. `http://creativecommons.org/licenses/by-sa/3.0/`

- cc-by-sa-2.5: Creative Commons Attribution ShareAlike 2.5 License. `http://creativecommons.org/licenses/by-sa/2.5/`

- cc-by-sa-2.0: Creative Commons Attribution ShareAlike 2.0 License. `http://creativecommons.org/licenses/by-sa/2.0/`

- cc-by-sa-1.0: Creative Commons Attribution ShareAlike 1.0 License. `http://creativecommons.org/licenses/by-sa/1.0/`

- cc-by-2.0: Creative Commons Attribution 2.0 License. `http://creativecommons.org/licenses/by/2.0/`

- cc-by-2.0: Creative Commons Attribution 2.0 License. `http://creativecommons.org/licenses/by/2.0/deed.en`

- cc-by-2.5: Creative Commons Attribution 2.5 License. `http://creativecommons.org/licenses/by/2.5/deed.en`

- cc-by-3.0: Creative Commons Attribution 3.0 License. `http://creativecommons.org/licenses/by/3.0/deed.en`

- GPL: GNU General Public License. `http://www.gnu.org/licenses/gpl-2.0.txt`

- LGPL: GNU Lesser General Public License. `http://www.gnu.org/licenses/lgpl.html`

- PD: This image is in the public domain.

- ATTR: The copyright holder of this file allows anyone to use it for any purpose, provided that the copyright holder is properly attributed. Redistribution, derivative work, commercial use, and all other use is permitted.

- EURO: This is the common (reverse) face of a euro coin. The copyright on the design of the common face of the euro coins belongs to the European Commission. Authorised is reproduction in a format without relief (drawings, paintings, films) provided they are not detrimental to the image of the euro.

- LFK: Lizenz Freie Kunst. `http://artlibre.org/licence/lal/de`

- CFR: Copyright free use.

- EPL: Eclipse Public License. `http://www.eclipse.org/org/documents/epl-v10.php`

Copies of the GPL, the LGPL as well as a GFDL are included in chapter Licenses[2]. Please note that images in the public domain do not require attribution. You may click on the image numbers in the following table to open the webpage of the images in your webbrower.

---

2   Chapter 9 on page 65

9 Licenses

9.1 GNU GENERAL PUBLIC LICENSE

Version 3, 29 June 2007

Copyright © 2007 Free Software Foundation, Inc. <http://fsf.org/>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed. Preamble

The GNU General Public License is a free, copyleft license for software and other kinds of works.

The licenses for most software and other practical works are designed to take away your freedom to share and change the works. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change all versions of a program–to make sure it remains free software for all its users. We, the Free Software Foundation, use the GNU General Public License for most of our software; it applies also to any other work released this way by its authors. You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for them if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs, and that you know you can do these things.

To protect your rights, we need to prevent others from denying you these rights or asking you to surrender the rights. Therefore, you have certain responsibilities if you distribute copies of the software, or if you modify it: responsibilities to respect the freedom of others.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must pass on to the recipients the same freedoms that you received. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

Developers that use the GNU GPL protect your rights with two steps: (1) assert copyright on the software, and (2) offer you this License giving you legal permission to copy, distribute and/or modify it.

For the developers' and authors' protection, the GPL clearly explains that there is no warranty for this free software. For both users' and authors' sake, the GPL requires that modified versions be marked as changed, so that their problems will not be attributed erroneously to authors of previous versions.

Some devices are designed to deny users access to install or run modified versions of the software inside them, although the manufacturer can do so. This is fundamentally incompatible with the aim of protecting users' freedom to change the software. The systematic pattern of such abuse occurs in the area of products for individuals to use, which is precisely where it is most unacceptable. Therefore, we have designed this version of the GPL to prohibit the practice for those products. If such problems arise substantially in other domains, we stand ready to extend this provision to those domains in future versions of the GPL, as needed to protect the freedom of users.

Finally, every program is threatened constantly by software patents. States should not allow patents to restrict development and use of software on general-purpose computers, but in those that do, we wish to avoid the special danger that patents applied to a free program could make it effectively proprietary. To prevent this, the GPL assures that patents cannot be used to render the program non-free.

The precise terms and conditions for copying, distribution and modification follow. TERMS AND CONDITIONS 0. Definitions.

"This License" refers to version 3 of the GNU General Public License.

"Copyright" also means copyright-like laws that apply to other kinds of works, such as semiconductor masks.

"The Program" refers to any copyrightable work licensed under this License. Each licensee is addressed as "you". "Licensees" and "recipients" may be individuals or organizations.

To "modify" a work means to copy from or adapt all or part of the work in a fashion requiring copyright permission, other than the making of an exact copy. The resulting work is called a "modified version" of the earlier work or a work "based on" the earlier work.

A "covered work" means either the unmodified Program or a work based on the Program.

To "propagate" a work means to do anything with it that, without permission, would make you directly or secondarily liable for infringement under applicable copyright law, except executing it on a computer or modifying a private copy. Propagation includes copying, distribution (with or without modification), making available to the public, and in some countries other activities as well.

To "convey" a work means any kind of propagation that enables other parties to make or receive copies. Mere interaction with a user through a computer network, with no transfer of a copy, is not conveying.

An interactive user interface displays "Appropriate Legal Notices" to the extent that it includes a convenient and prominently visible feature that (1) displays an appropriate copyright notice, and (2) tells the user that there is no warranty for the work (except to the extent that warranties are provided), that licensees may convey the work under this License, and how to view a copy of this License. If the interface presents a list of user commands or options, such as a menu, a prominent item in the list meets this criterion. 1. Source Code.

The "source code" for a work means the preferred form of the work for making modifications to it. "Object code" means any non-source form of a work.

A "Standard Interface" means an interface that either is an official standard defined by a recognized standards body, or, in the case of interfaces specified for a particular programming language, one that is widely used among developers working in that language.

The "System Libraries" of an executable work include anything, other than the work as a whole, that (a) is included in the normal form of packaging a Major Component, but which is not part of that Major Component, and (b) serves only to enable use of the work with that Major Component, or to implement a Standard Interface for which an implementation is available to the public in source code form. A "Major Component", in this context, means a major essential component (kernel, window system, and so on) of the specific operating system (if any) on which the executable work runs, or a compiler used to produce the work, or an object code interpreter used to run it.

The "Corresponding Source" for a work in object code form means all the source code needed to generate, install, and (for an executable work) run the object code and to modify the work, including scripts to control those activities. However, it does not include the work's System Libraries, or general-purpose tools or generally available free programs which are used unmodified in performing those activities but which are not part of the work. For example, Corresponding Source includes interface definition files associated with source files for the work, and the source code for shared libraries and dynamically linked subprograms that the work is specifically designed to require, such as by intimate data communication or control flow between those subprograms and other parts of the work.

The Corresponding Source need not include anything that users can regenerate automatically from other parts of the Corresponding Source.

The Corresponding Source for a work in source code form is that same work. 2. Basic Permissions.

All rights granted under this License are granted for the term of copyright on the Program, and are irrevocable provided the stated conditions are met. This License explicitly affirms your unlimited permission to run the unmodified Program. The output from running a covered work is covered by this License only if the output, given its content, constitutes a covered work. This License acknowledges your rights of fair use or other equivalent, as provided by copyright law.

You may make, run and propagate covered works that you do not convey, without conditions so long as your license otherwise remains in force. You may convey covered works to others for the sole purpose of having them make modifications exclusively for you, or provide you with facilities for running those works, provided that you comply with the terms of this License in conveying all material for which you do not control copyright. Those thus making or running the covered works for you must do so exclusively on your behalf, under your direction and control, on terms that prohibit them from making any copies of your copyrighted material outside their relationship with you.

Conveying under any other circumstances is permitted solely under the conditions stated below. Sublicensing is not allowed; section 10 makes it unnecessary. 3. Protecting Users' Legal Rights From Anti-Circumvention Law.

No covered work shall be deemed part of an effective technological measure under any applicable law fulfilling obligations under article 11 of the WIPO copyright treaty adopted on 20 December 1996, or similar laws prohibiting or restricting circumvention of such measures.

When you convey a covered work, you waive any legal power to forbid circumvention of technological measures to the extent such circumvention is effected by exercising rights under this License with respect to the covered work, and you disclaim any intention to limit operation or modification of the work as a means of enforcing, against the work's users, your or third parties' legal rights to forbid circumvention of technological measures. 4. Conveying Verbatim Copies.

You may convey verbatim copies of the Program's source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice; keep intact all notices stating that this License and any non-permissive terms added in accord with section 7 apply to the code; keep intact all notices of the absence of any warranty; and give all recipients a copy of this License along with the Program.

You may charge any price or no price for each copy that you convey, and you may offer support or warranty protection for a fee. 5. Conveying Modified Source Versions.

You may convey a work based on the Program, or the modifications to produce it from the Program, in the form of source code under the terms of section 4, provided that you also meet all of these conditions:

* a) The work must carry prominent notices stating that you modified it, and giving a relevant date. * b) The work must carry prominent notices stating that it is released under this License and any conditions added under section 7. This requirement modifies the requirement in section 4 to "keep intact all notices". * c) You must license the entire work, as a whole, under this License to anyone who comes into possession of a copy. This License will therefore apply, along with any applicable section 7 additional terms, to the whole of the work, and all its parts, regardless of how they are packaged. This License gives no permission to license the work in any other way, but it does not invalidate such permission if you have separately received it. * d) If the work has interactive user interfaces, each must display Appropriate Legal Notices; however, if the Program has interactive interfaces that do not display Appropriate Legal Notices, your work need not make them do so.

A compilation of a covered work with other separate and independent works, which are not by their nature extensions of the covered work, and which are not combined with it such as to form a larger program, in or on a volume of a storage or distribution medium, is called an "aggregate" if the compilation and its resulting copyright are not used to limit the access or legal rights of the compilation's users beyond what the individual works permit. Inclusion of a covered work in an aggregate does not cause this License to apply to the other parts of the aggregate. 6. Conveying Non-Source Forms.

You may convey a covered work in object code form under the terms of sections 4 and 5, provided that you also convey the machine-readable Corresponding Source under the terms of this License, in one of these ways:

* a) Convey the object code in, or embodied in, a physical product (including a physical distribution medium), accompanied by the Corresponding Source fixed on a durable physical medium customarily used for software interchange. * b) Convey the object code in, or embodied in, a physical product (including a physical distribution medium), accompanied by a written offer, valid for at least three years and valid for as long as you offer spare parts or customer support for that product model, to give anyone who possesses the object code either (1) a copy of the Corresponding Source for all the software in the product that is covered by this License, on a durable physical medium customarily used for software interchange, for a price no more than your reasonable cost of physically performing this conveying of source, or (2) access to copy the Corresponding Source from a network server at no charge. * c) Convey individual copies of the object code with a copy of the written offer to provide the Corresponding Source. This alternative is allowed only occasionally and noncommercially, and only if you received the object code with such an offer, in accord with subsection 6b. * d) Convey the object code by offering access from a designated place (gratis or for a charge), and offer equivalent access to the Corresponding Source in the same way through the same place at no further charge. You need not require recipients to copy the Corresponding Source along with the object code. If the place to copy the object code is a network server, the Corresponding Source may be on a different server (operated by you or a third party) that supports equivalent copying facilities, provided you maintain clear directions next to the object code saying where to find the Corresponding Source. Regardless of what server hosts the Corresponding Source, you remain obligated to ensure that it is available for as long as needed to satisfy these requirements. * e) Convey the object code using peer-to-peer transmission, provided you inform other peers where the object code and Corresponding Source of the work are being offered to the general public at no charge under subsection 6d.

A separable portion of the object code, whose source code is excluded from the Corresponding Source as a System Library, need not be included in conveying the object code work.

A "User Product" is either (1) a "consumer product", which means any tangible personal property which is normally used for personal, family, or household purposes, or (2) anything designed or sold for incorporation into a dwelling. In determining whether a product is a consumer product, doubtful cases shall be resolved in favor of coverage. For a particular product received by a particular user, "normally used" refers to a typical or common use of that class of product, regardless of the status of the particular user or of the way in which the particular user actually uses, or expects or is expected to use, the product. A product is a consumer product regardless of whether the product has substantial commercial, industrial or non-consumer uses, unless such uses represent the only significant mode of use of the product.

"Installation Information" for a User Product means any methods, procedures, authorization keys, or other information required to install and execute modified versions of a covered work in that User Product from a modified version of its Corresponding Source. The information must suffice to ensure that the continued functioning of the modified object code is in no case prevented or interfered with solely because modification has been made.

If you convey an object code work under this section in, or with, or specifically for use in, a User Product, and the conveying occurs as part of a transaction in which the right of possession and use of the User Product is transferred to the recipient in perpetuity or for a fixed term (regardless of how the transaction is characterized), the Corresponding Source conveyed under this section must be accompanied by the Installation Information. But this requirement does not apply if neither you nor any third party retains the ability to install modified object code on the User Product (for example, the work has been installed in ROM).

The requirement to provide Installation Information does not include a requirement to continue to provide support service, warranty, or updates for a work that has been modified or installed by the recipient, or for the User Product in which it has been modified or installed. Access to a network may be denied when the modification itself materially and adversely affects the operation of the network or violates the rules and protocols for communication across the network.

Corresponding Source conveyed, and Installation Information provided, in accord with this section must be in a format that is publicly documented (and with an implementation available to the public in source code form), and must require no special password or key for unpacking, reading or copying. 7. Additional Terms.

"Additional permissions" are terms that supplement the terms of this License by making exceptions from one or more of its conditions. Additional permissions that are applicable to the entire Program shall be treated as though they were included in this License, to the extent that they are valid under applicable law. If additional permissions apply only to part of the Program, that part may be used separately under those permissions, but the entire Program remains governed by this License without regard to the additional permissions.

When you convey a copy of a covered work, you may at your option remove any additional permissions from that copy, or from any part of it. (Additional permissions may be written to require their own removal in certain cases when you modify the work.) You may place additional permissions on material, added by you to a covered work, for which you have or can give appropriate copyright permission.

Notwithstanding any other provision of this License, for material you add to a covered work, you may (if authorized by the copyright holders of that material) supplement the terms of this License with terms:

* a) Disclaiming warranty or limiting liability differently from the terms of sections 15 and 16 of this License; or * b) Requiring preservation of specified reasonable legal notices or author attributions in that material or in the Appropriate Legal Notices displayed by works containing it; or * c) Prohibiting misrepresentation of the origin of that material, or requiring that modified versions of such material be marked in reasonable ways as different from the original version; or * d) Limiting the use for publicity purposes of names of licensors or authors of the material; or * e) Declining to grant rights under trademark law for use of some trade names, trademarks, or service marks; or * f) Requiring indemnification of licensors and authors of that material by anyone who conveys the material (or modified versions of it) with contractual assumptions of liability to the recipient, for any liability that these contractual assumptions directly impose on those licensors and authors.

All other non-permissive additional terms are considered "further restrictions" within the meaning of section 10. If the Program as you received it, or any part of it, contains a notice stating that it is governed by this License along with a term that is a further restriction, you may remove that term. If a license document contains a further restriction but permits relicensing or conveying under this License, you may add to a covered work material governed by the terms of that license document, provided that the further restriction does not survive such relicensing or conveying.

If you add terms to a covered work in accord with this section, you must place, in the relevant source files, a statement of the additional terms that apply to those files, or a notice indicating where to find the applicable terms.

Additional terms, permissive or non-permissive, may be stated in the form of a separately written license, or stated as exceptions; the above requirements apply either way. 8. Termination.

You may not propagate or modify a covered work except as expressly provided under this License. Any attempt otherwise to propagate or modify it is void, and will automatically terminate your rights under this License (including any patent licenses granted under the third paragraph of section 11).

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, you do not qualify to receive new licenses for the same material under section 10. 9. Acceptance Not Required for Having Copies.

You are not required to accept this License in order to receive or run a copy of the Program. Ancillary propagation of a covered work occurring solely as a consequence of using peer-to-peer transmission to receive a copy likewise does not require acceptance. However, nothing other than this License grants you permission to propagate or modify any covered work. These actions infringe copyright if you do not accept this License. Therefore, by modifying or propagating a covered work, you indicate your acceptance of this License to do so. 10. Automatic Licensing of Downstream Recipients.

Each time you convey a covered work, the recipient automatically receives a license from the original licensors, to run, modify and propagate that work, subject to this License. You are not responsible for enforcing compliance by third parties with this License.

An "entity transaction" is a transaction transferring control of an organization, or substantially all assets of one, or subdividing an organization, or merging organizations. If propagation of a covered work results from an entity transaction, each party to that transaction who receives a copy of the work also receives whatever licenses to the work the party's predecessor in interest had or could give under the previous paragraph, plus a right to possession of the Corresponding Source of the work from the predecessor in interest, if the predecessor has it or can get it with reasonable efforts.

You may not impose any further restrictions on the exercise of the rights granted or affirmed under this License. For example, you may not impose a license fee, royalty, or other charge for exercise of rights granted under this License, and you may not initiate litigation (including a cross-claim or counterclaim in a lawsuit) alleging that any patent claim is infringed by making, using, selling, offering for sale, or importing the Program or any portion of it. 11. Patents.

A "contributor" is a copyright holder who authorizes use under this License of the Program or a work on which the Program is based. The work thus licensed is called the contributor's "contributor version".

A contributor's "essential patent claims" are all patent claims owned or controlled by the contributor, whether already acquired or hereafter acquired, that would be infringed by some manner, permitted by this License, of making, using, or selling its contributor version, but do not include claims that would be infringed only as a consequence of further modification of the contributor version. For purposes of this definition, "control" includes the right to grant patent sublicenses in a manner consistent with the requirements of this License.

Each contributor grants you a non-exclusive, worldwide, royalty-free patent license under the contributor's essential patent claims, to make, use, sell, offer for sale, import and otherwise run, modify and propagate the contents of its contributor version.

In the following three paragraphs, a "patent license" is any express agreement or commitment, however denominated, not to enforce a patent (such as an express permission to practice a patent or covenant not to sue for patent infringement). To "grant" such a patent license to a party means to make such an agreement or commitment not to enforce a patent against the party.

If you convey a covered work, knowingly relying on a patent license, and the Corresponding Source of the work is not available for anyone to copy, free of charge and under the terms of this License, through a publicly available network server or other readily accessible means, then you must either (1) cause the Corresponding Source to be so available, or (2) arrange to deprive yourself of the benefit of the patent license for this particular work, or (3) arrange, in a manner consistent with the requirements of this License, to extend the patent license to downstream recipients. "Knowingly relying" means you have actual knowledge that, but for the patent license, your conveying the covered work in a country, or your recipient's use of the covered work in a country, would infringe one or more identifiable patents in that country that you have reason to believe are valid.

If, pursuant to or in connection with a single transaction or arrangement, you convey, or propagate by procuring conveyance of, a covered work, and grant a patent license to some of the parties receiving the covered work authorizing them to use, propagate, modify or convey a specific copy of the covered work, then the patent license you grant is automatically extended to all recipients of the covered work and works based on it.

A patent license is "discriminatory" if it does not include within the scope of its coverage, prohibits the exercise of, or is conditioned on the non-exercise of one or more of the rights that are specifically granted under this License. You may not convey a covered work if you are a party to an arrangement with a third party that is in the business of distributing software, under which you make payment to the third party based on the extent of your activity of conveying the work, and under which the third party grants, to any of the parties who would receive the covered work from you, a discriminatory patent license (a) in connection with copies of the covered work conveyed by you (or copies made from those copies), or (b) primarily for and in connection with specific products or compilations that contain the covered work, unless you entered into that arrangement, or that patent license was granted, prior to 28 March 2007.

Nothing in this License shall be construed as excluding or limiting any implied license or other defenses to infringement that may otherwise be available to you under applicable patent law. 12. No Surrender of Others' Freedom.

If conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot convey a covered work so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not convey it at all. For example, if you agree to terms that obligate you to collect a royalty for further conveying from those to whom you convey the Program, the only way you could satisfy

both those terms and this License would be to refrain entirely from conveying the Program. 13. Use with the GNU Affero General Public License.

Notwithstanding any other provision of this License, you have permission to link or combine any covered work with a work licensed under version 3 of the GNU Affero General Public License into a single combined work, and to convey the resulting work. The terms of this License will continue to apply to the part which is the covered work, but the special requirements of the GNU Affero General Public License, section 13, concerning interaction through a network will apply to the combination as such. 14. Revised Versions of this License.

The Free Software Foundation may publish revised and/or new versions of the GNU General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Program specifies that a certain numbered version of the GNU General Public License "or any later version" applies to it, you have the option of following the terms and conditions either of that numbered version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of the GNU General Public License, you may choose any version ever published by the Free Software Foundation.

If the Program specifies that a proxy can decide which future versions of the GNU General Public License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Program.

Later license versions may give you additional or different permissions. However, no additional obligations are imposed on any author or copyright holder as a result of your choosing to follow a later version. 15. Disclaimer of Warranty.

THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION. 16. Limitation of Liability.

IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MODIFIES AND/OR CONVEYS THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. 17. Interpretation of Sections 15 and 16.

If the disclaimer of warranty and limitation of liability provided above cannot be given local legal effect according to their terms, reviewing courts shall apply local law that most closely approximates an absolute waiver of all civil liability in connection with the Program, unless a warranty or assumption of liability accompanies a copy of the Program in return for a fee.

END OF TERMS AND CONDITIONS How to Apply These Terms to Your New Programs

If you develop a new program, and you want it to be of the greatest possible use to the public, the best way to achieve this is to make it free software which everyone can redistribute and change under these terms.

To do so, attach the following notices to the program. It is safest to attach them to the start of each source file to most effectively state the exclusion of warranty; and each file should have at least the "copyright" line and a pointer to where the full notice is found.

<one line to give the program's name and a brief idea of what it does.> Copyright (C) <year> <name of author>

This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program. If not, see <http://www.gnu.org/licenses/>.

Also add information on how to contact you by electronic and paper mail.

If the program does terminal interaction, make it output a short notice like this when it starts in an interactive mode:

<program> Copyright (C) <year> <name of author> This program comes with ABSOLUTELY NO WARRANTY; for details type 'show w'. This is free software, and you are welcome to redistribute it under certain conditions; type 'show c' for details.

The hypothetical commands 'show w' and 'show c' should show the appropriate parts of the General Public License. Of course, your program's commands might be different; for a GUI interface, you would use an "about box".

You should also get your employer (if you work as a programmer) or school, if any, to sign a "copyright disclaimer" for the program, if necessary. For more information on this, and how to apply and follow the GNU GPL, see <http://www.gnu.org/licenses/>.

The GNU General Public License does not permit incorporating your program into proprietary programs. If your program is a subroutine library, you may consider it more useful to permit linking proprietary applications with the library. If this is what you want to do, use the GNU Lesser General Public License instead of this License. But first, please read <http://www.gnu.org/philosophy/why-not-lgpl.html>.

9.2 GNU Free Documentation License

Version 1.3, 3 November 2008

Copyright © 2000, 2001, 2002, 2007, 2008 Free Software Foundation, Inc. <http://fsf.org/>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed. 0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference. 1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person or entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses

following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License. 2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies. 3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document. 4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

* A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission. * B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement. * C. State on the Title page the name of the publisher of the Modified Version, as the publisher. * D. Preserve all the copyright notices of the Document. * E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices. * F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below. * G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice. * H. Include an unaltered copy of this License. * I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence. * J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission. * K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein. * L. Preserve all the Invariant Sections of the Document, unaltered in their text and

in their titles. Section numbers or the equivalent are not considered part of the section titles. * M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version. * N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section. * O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version. 5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements". 6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document. 7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate. 8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title

(section 1) will typically require changing the actual title. 9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it. 10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document. 11. RELICENSING

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing. ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

Copyright (C) YEAR YOUR NAME. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with ... Texts." line with this:

with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

9.3 GNU Lesser General Public License

GNU LESSER GENERAL PUBLIC LICENSE

Version 3, 29 June 2007

Copyright © 2007 Free Software Foundation, Inc. <http://fsf.org/>

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

This version of the GNU Lesser General Public License incorporates the terms and conditions of version 3 of the GNU General Public License, supplemented by the additional permissions listed below. 0. Additional Definitions.

As used herein, "this License" refers to version 3 of the GNU Lesser General Public License, and the "GNU GPL" refers to version 3 of the GNU General Public License.

"The Library" refers to a covered work governed by this License, other than an Application or a Combined Work as defined below.

An "Application" is any work that makes use of an interface provided by the Library, but which is not otherwise based on the Library. Defining a subclass of a class defined by the Library is deemed a mode of using an interface provided by the Library.

A "Combined Work" is a work produced by combining or linking an Application with the Library. The particular version of the Library with which the Combined Work was made is also called the "Linked Version".

The "Minimal Corresponding Source" for a Combined Work means the Corresponding Source for the Combined Work, excluding any source code for portions of the Combined Work that, considered in isolation, are based on the Application, and not on the Linked Version.

The "Corresponding Application Code" for a Combined Work means the object code and/or source code for the Application, including any data and utility programs needed for reproducing the Combined Work from the Application, but excluding the System Libraries of the Combined Work. 1. Exception to Section 3 of the GNU GPL.

You may convey a covered work under sections 3 and 4 of this License without being bound by section 3 of the GNU GPL. 2. Conveying Modified Versions.

If you modify a copy of the Library, and, in your modifications, a facility refers to a function or data to be supplied by an Application that uses the facility (other than as an argument passed when the facility is invoked), then you may convey a copy of the modified version:

* a) under this License, provided that you make a good faith effort to ensure that, in the event an Application does not supply the function or data, the facility still operates, and performs whatever part of its purpose remains meaningful, or * b) under the GNU GPL, with none of the additional permissions of this License applicable to that copy.

3. Object Code Incorporating Material from Library Header Files.

The object code form of an Application may incorporate material from a header file that is part of the Library. You may convey such object code under terms of your choice, provided that, if the incorporated material is not limited to numerical parameters, data structure layouts and accessors, or small macros, inline functions and templates (ten or fewer lines in length), you do both of the following:

* a) Give prominent notice with each copy of the object code that the Library is used in it and that the Library and its use are covered by this License. * b) Accompany the object code with a copy of the GNU GPL and this license document.

4. Combined Works.

You may convey a Combined Work under terms of your choice that, taken together, effectively do not restrict modification of the portions of the Library contained in the Combined Work and reverse engineering for debugging such modifications, if you also do each of the following:

* a) Give prominent notice with each copy of the Combined Work that the Library is used in it and that the Library and its use are covered by this License. * b) Accompany the Combined Work with a copy of the GNU GPL and this license document. * c) For a Combined Work that displays copyright notices during execution, include the copyright notice for the Library among these notices, as well as a reference directing the user to the copies of the GNU GPL and this license document. * d) Do one of the following: o 0) Convey the Minimal Corresponding Source under the terms of this License, and the Corresponding Application Code in a form suitable for, and under terms that permit, the user to recombine or relink the Application with a modified version of the Linked Version to produce a modified Combined Work, in the manner specified by section 6 of the GNU GPL for conveying Corresponding Source. o 1) Use a suitable shared library mechanism for linking with the Library. A suitable mechanism is one that (a) uses at run time a copy of the Library already present on the user's computer system, and (b) will operate properly with a modified version of the Library that is interface-compatible with the Linked Version. * e) Provide Installation Information, but only if you would otherwise be required to provide such information under section 6 of the GNU GPL, and only to the extent that such information is necessary to install and execute a modified version of the Combined Work produced by recombining or relinking the Application with a modified version of the Linked Version. (If you use option 4d0, the Installation Information must accompany the Minimal Corresponding Source and Corresponding Application Code. If you use option 4d1, you must provide the Installation Information in the manner specified by section 6 of the GNU GPL for conveying Corresponding Source.)

5. Combined Libraries.

You may place library facilities that are a work based on the Library side by side in a single library together with other library facilities that are not Applications and are not covered by this License, and convey such a combined library under terms of your choice, if you do both of the following:

* a) Accompany the combined library with a copy of the same work based on the Library, uncombined with any other library facilities, conveyed under the terms of this License. * b) Give prominent notice with the combined library that part of it is a work based on the Library, and explaining where to find the accompanying uncombined form of the same work.

6. Revised Versions of the GNU Lesser General Public License.

The Free Software Foundation may publish revised and/or new versions of the GNU Lesser General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

Each version is given a distinguishing version number. If the Library as you received it specifies that a certain numbered version of the GNU Lesser General Public License "or any later version" applies to it, you have the option of following the terms and conditions either of that published version or of any later version published by the Free Software Foundation. If the Library as you received it does not specify a version number of the GNU Lesser General Public License, you may choose any version of the GNU Lesser General Public License ever published by the Free Software Foundation.

If the Library as you received it specifies that a proxy can decide whether future versions of the GNU Lesser General Public License shall apply, that proxy's public statement of acceptance of any version is permanent authorization for you to choose that version for the Library.