

事業計画及び成長可能性に関する説明資料

【 AEIとChatGPTの関係性に関する資料の補足資料 】

株式会社 **pluszero** (証券コード: 5132)

2023年4月13日

資料の取り扱いについて

本資料については、技術的な内容を含んでおり、解説動画も併せて開示しておりますので、ご参考いただけますと幸いです。

<https://aa>

本資料には、将来の見通しに関する記述が含まれています。これらの将来の見通しに関する記述は、本資料の日付時点の情報に基づいて作成されています。これらの記述は、将来の結果や業績を保証するものではありません。このような将来予想に関する記述には、必ずしも既知および未知のリスクや不確実性が含まれており、その結果、将来の実際の業績や財務状況は、将来予想に関する記述によって明示的または黙示的に示された将来の業績や結果の予測とは大きく異なる可能性があります。

これらの記述に記載された結果と大きく異なる可能性のある要因には、国内および国際的な経済状況の変化や、当社が事業を展開する業界の動向などが含まれますが、これらに限定されるものではありません。

また、当社以外の事項・組織に関する情報は、一般に公開されている情報に基づいており、当社はそのような一般に公開されている情報の正確性や適切性を検証しておらず、保証しておりません。

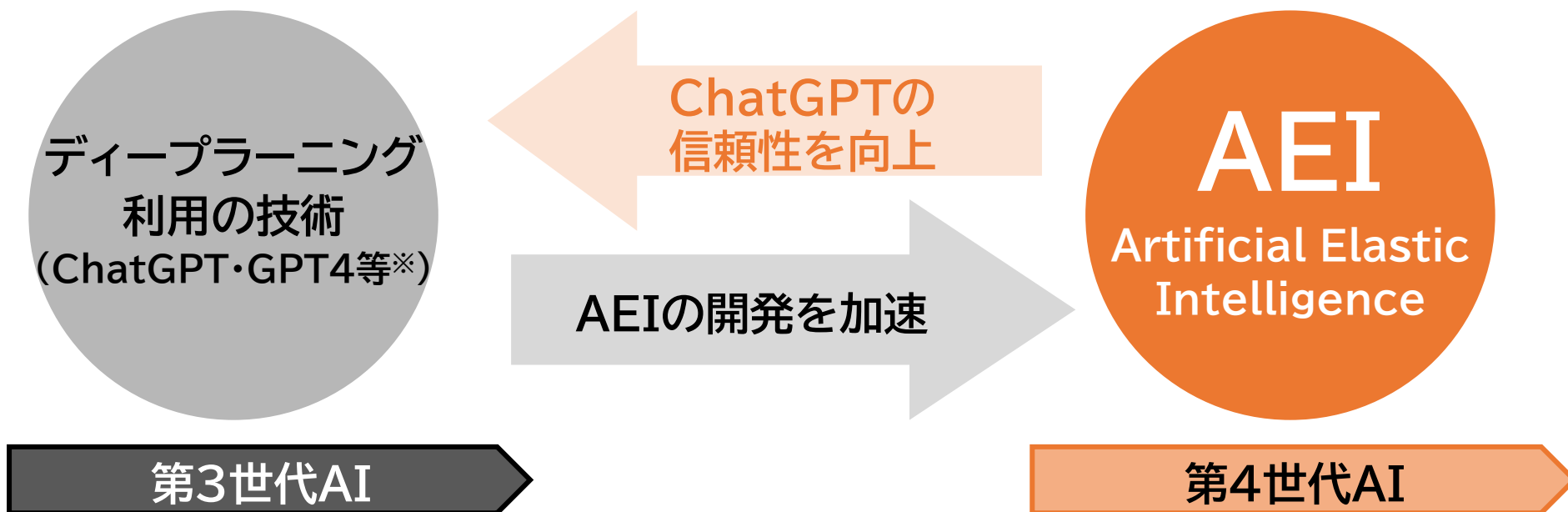
AEIとChatGPTの関係性

ChatGPTはAEIにとって強烈な追い風になる

I ChatGPTとAEIは相互補完関係

- ① AEIはChatGPTの信頼性を向上させる
- ② ChatGPTはAEIの開発を加速させる

II ChatGPTは、AEIが戦う「言語系AI」の市場の可能性の高さを世の中に認知^{*1*2}させた



*1: Microsoft社による10億ドルの出資(Bloomberg)

*2: 文科省による国内での教育現場への活用の検討(日本経済新聞)

※: ChatGPTはサービス名、GPT4はChatGPTで使用されている大規模言語モデルになります。本資料においてChatGPTは大規模言語モデル含んだ類似のサービス全てを指します。

AEIとChatGPTの比較

		ChatGPT	AEI	AEI+ChatGPT
AIの対応範囲		広範		業界特化 *業界拡大は26年10月期以降加速
AIの性能	精度	○	△→◎ *23/10期～26/10期に向上	○→◎ *23/10期～26/10期に向上
	信頼性	△→? *ディープラーニングの外部機能次第		◎
	現場適用力	△→? *現場依存・プラグインの進化次第		◎
AIの開発難易度	費用面	△	○	◎ *知能の整備効率向上
	人材面	理系人材中心 *高等数学、計算機科学など		文理融合人材中心
	技術面	大規模言語モデル	意味の見える化	意味の見える化 +大規模言語モデル

pluszeroが定義するAIの性能評価軸

- 当社はAIの性能評価の軸として、**精度と信頼性**の2軸があると考えている。

① 精度

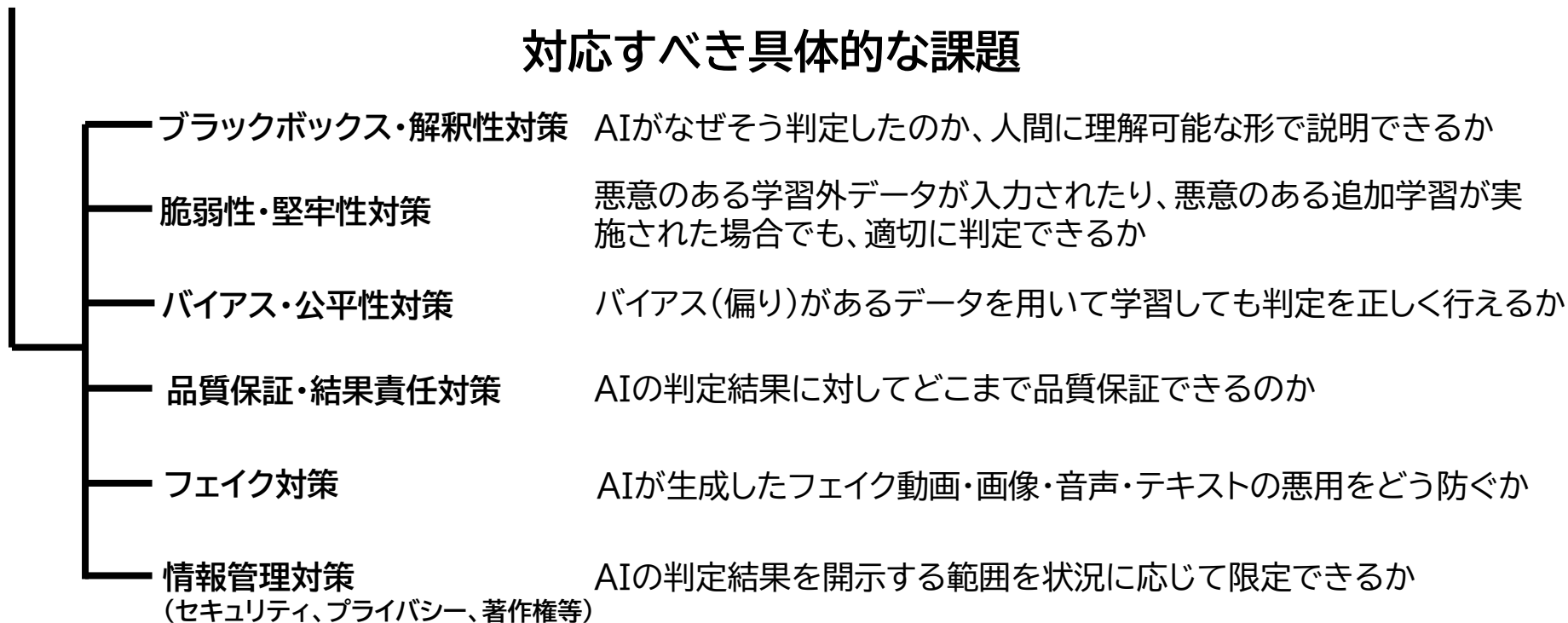
AIの出した答えがどれだけ正解かを示した指標

② 信頼性

AIの出した答えをどれだけ信用できるかを示した指標

- 具体的な課題に対しての対応能力に分解することが可能
- 文科省系研究機関であるCRDS定義の「**信頼されるAI**」に示されている5つの要素
- セキュリティ/プライバシー/著作権等の情報管理対策

対応すべき具体的な課題



ChatGPTは穴埋め問題のスペシャリスト

- ある文章があったときに、次に続く単語(厳密にはトークン)を誰よりも上手く予測するもの。
- 例えて言うと、文章の末尾の穴を埋める穴埋め問題のスペシャリスト

穴埋め問題のスペシャリスト

たまに間違える物知り博士

事実でも架空の話でも、とにかく穴をそれらしく埋めることができる。穴埋めをするときに、統計的・確率的観点で尤もらしいもので埋める。ある意味、たまに間違える物知り博士といえる。

何でも答えられる

精度
◎

あらゆる話題に対して自然な回答ができる*3。

思考回路が読めない

信頼性
(説明可能性)
△

自己教師あり学習を中心としたディープラーニングベースの技術である。だから、「ブラックボックスで中身が見えない」という問題がある*4。

自信満々に間違える

信頼性
(品質保証)
△

答えがないものや曖昧なものにも適当に答えてしまう場合があるので、正確に答えることも多いが、時には「自信満々に間違える」傾向がある*5。

*3: 様々な試験問題や研究ベンチマークで高スコアを達成しています。(OpenAI公式ブログ)

*4: 入力クエリにstep by stepといれるとある程度の思考過程は見るすることができます。(Chain of thought prompting elicits reasoning in large language models)

*5: hallucinationの問題は解決していません。(OpenAI公式ブログ)

ChatGPTの信頼性向上の試みと依然として残る課題

ChatGPTは、信頼性を高めるための独自の取り組みを行っている。
しかし、ディープラーニングという根本の技術に起因して、「自信満々に間違える」傾向や「思考回路が読めない」傾向があるので、信頼性の解決の難易度は極めて高い。

ChatGPTの取り組み

人間視点で良い出力

人間にとって、より受け入れられやすい応答になるように学習している*6

有害なコンテンツの
フィルタリング

出力にフィルタをかけることによって、有害なコンテンツが含まれているかを評価して、アプリケーションの有害な動作を防ぐ*7*8*9

主な残課題

思考回路が読めない

自信満々に間違える

信頼性(説明可能性) △

信頼性(品質保証) △

*6: 様々な話題で人間の好む回答になるように学習しています。(Training language models to follow instructions with human feedback)

*7: フィルターは回避される場合があります。(GPT-4 Technical Report)

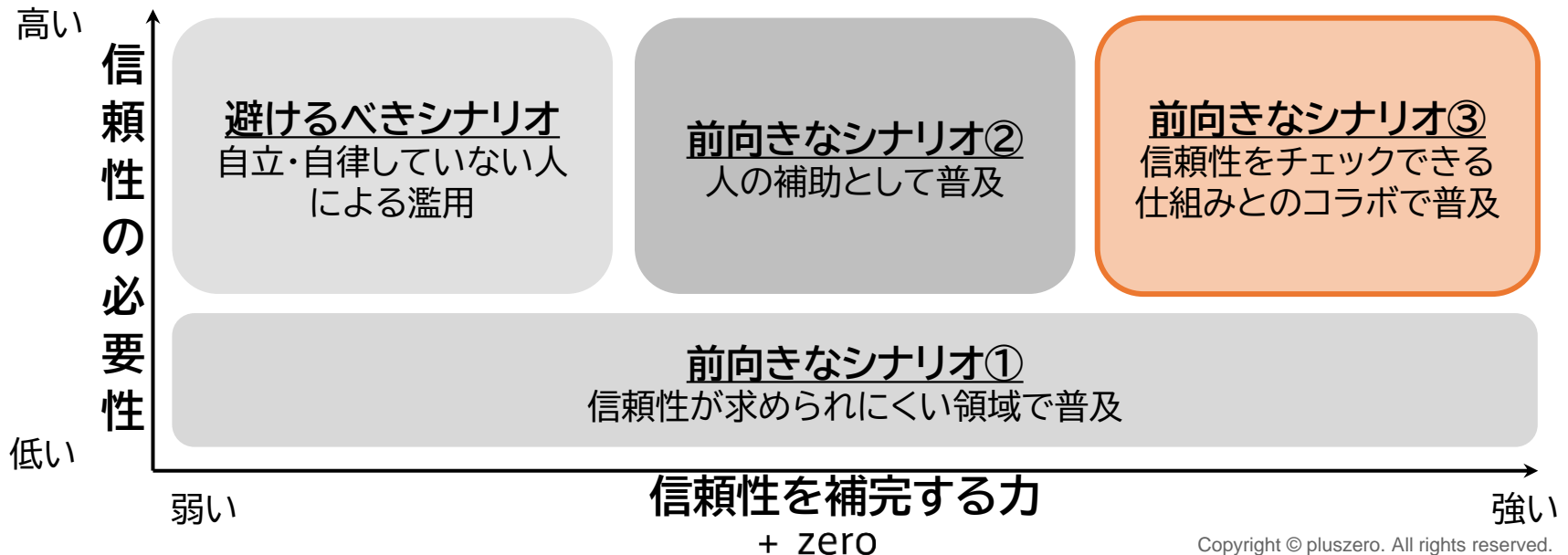
*8: クエリに有害なコンテンツが含まれているかどうかを検知する仕組みも含まれます。(OpenAI公式ブログ)

*9: 学習時に有害な出力を避けるように学習する工夫を実施しています。(GPT-4 System Card)

ChatGPTの未来は信頼性によって予測できる

ChatGPTの普及の道筋は信頼性が決める

- 前向きなシナリオ① 信頼性が求められにくい領域で普及する
エンタメなどのクリエイティブ領域や説明責任が問われにくい課題など
- 前向きなシナリオ② 信頼性を自らチェックできる人を補助する役割として普及する
検索、要約、プログラミングの補助、資料の叩き台、ブレストの相手など
- 前向きなシナリオ③ 信頼性を自動でチェックできる仕組み(AEIなど)と組み合わせることで普及する
信頼性のチェックも任せられれば、全体の自動化率が向上する
- 避けるべきシナリオ 信頼性を自らチェックできない人が結果を鵜呑みにして濫用する



信頼性が求められにくい領域 ～前向きなシナリオ①の補足説明～

① アウトプットが出ること自体に価値があるもの
(明確な精度の定義が難しいもの)

② 説明責任(解釈性)・結果責任が問われにくい課題

クリエイティブ領域

芸術・デザイン・エンターテインメント

↳ 動画・画像・音声・音楽・テキスト生成

アイデア

↳ ブレインストーミング

※ 著作権、プライバシーなどは考慮する必要がある

社会に悪影響が無い範囲で使う

個人で閉じる場合

↳ 個人の趣味嗜好の中で用いるなど

社会と接点がある場合*10

↳ 間違った答えが含まれることを織込済のケース
例：ネット検索など

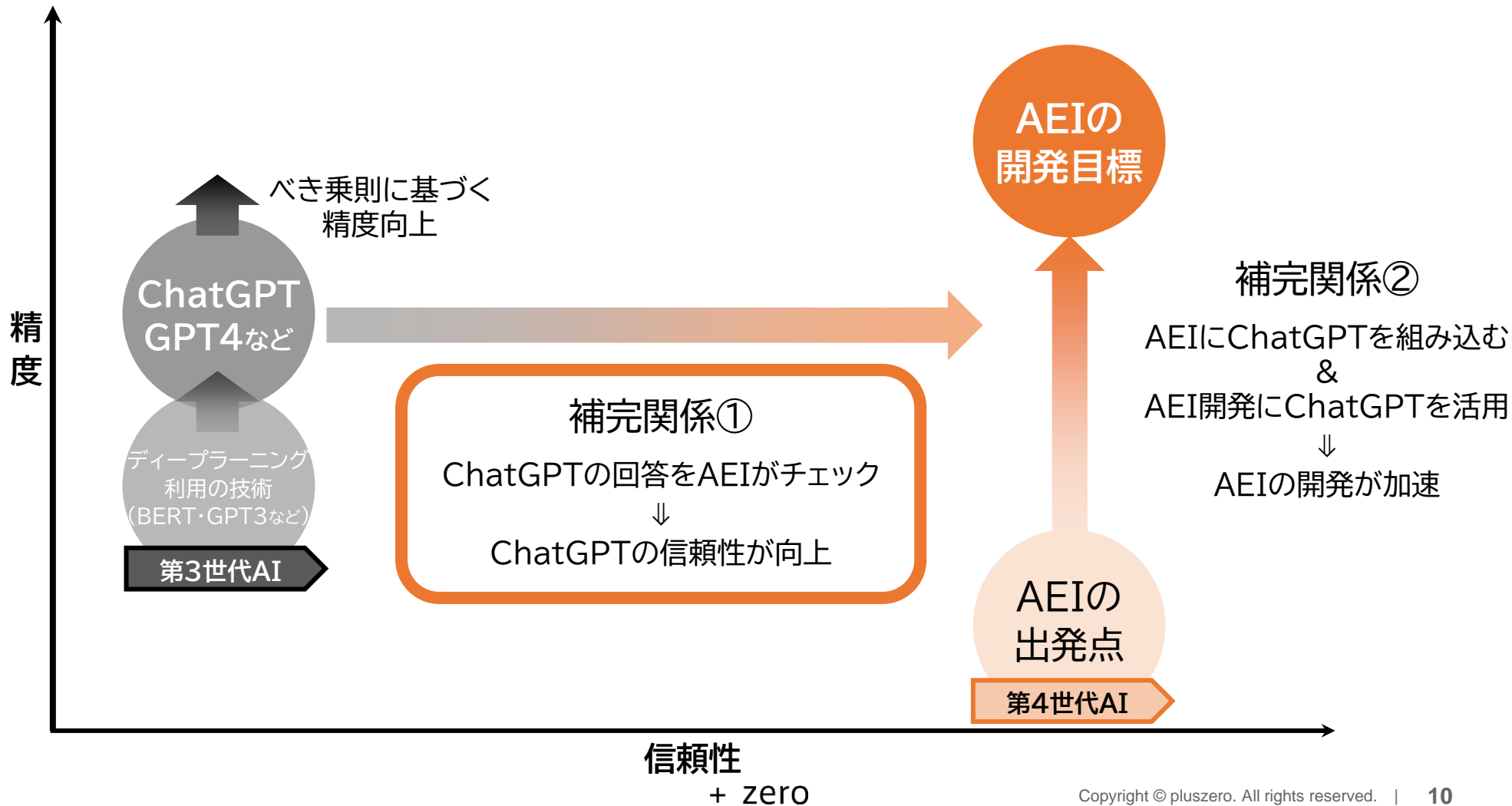
↳ 精度が一定以上に高ければ、解釈性は問われにくいケース
例：画像認識、OCRなど

*10: ChatGPTの利用規約は一部上記の内容を踏まえて設定されています。

※: ①②以外にも、ある分野においてChatGPTの精度がほぼ100%となった場合は、その精度自体が信頼性を生み、ChatGPT単体で利用できる可能性があります。

AEIはChatGPTの信頼性を向上させる

- ChatGPTの回答をAEIがチェック
- ChatGPTの信頼性を高めることができる



信頼性を高めるために必要なこと

信頼性に関する具体的な課題を解決するためには、
「見える化」と「ルールとの適合性」と「認証・アクセス制御」が有効であり、AEIで対策可能である。

具体的な課題

ブラックボックス・解釈可能性対策

脆弱性・堅牢性対策

バイアス・公平性対策

品質保証・結果責任対策

フェイク対策

セキュリティー対策

プライバシー対策

著作権対策

見える化

ルールとの適合性
(決まり・規則・法則に適しているか)

認証
(本人・本物確認)
アクセス制御
(誰に何を出すか)

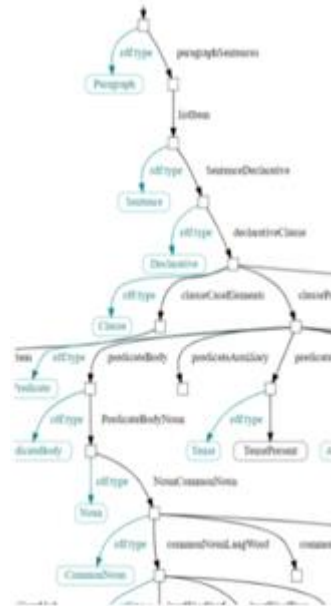
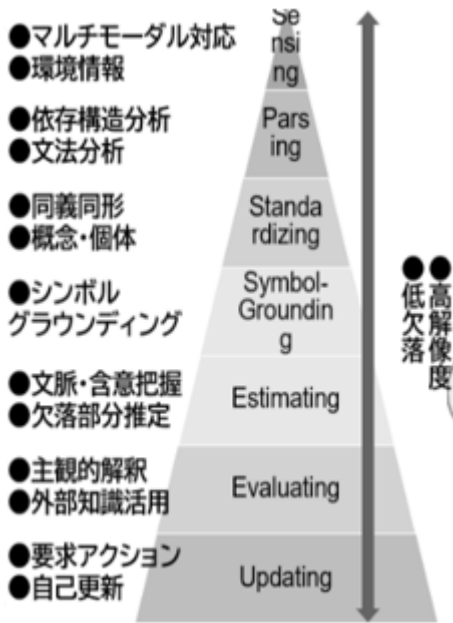
AEIで対策可能

AEIは見える化を追求した技術である

意味と知能の見える化

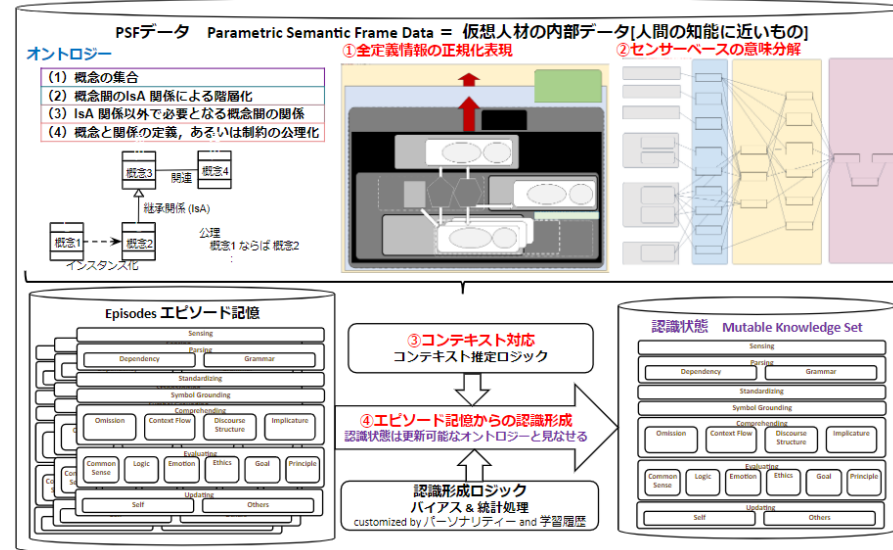
意味の見える化

言葉が含む意味を機械(AI)が扱う形式にする
→ N4で実現



知能の見える化

一人一人の頭の中の違いをモデル化する
→ PSFデータで実現



AEIは「見える化」により「ルールとの適合性」を実現する

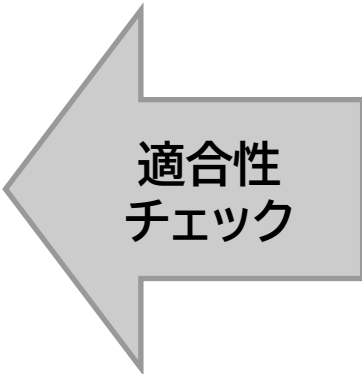
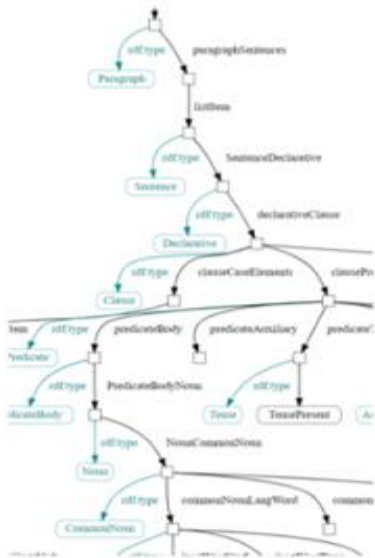
見える化したからこそ、ルールとの適合性をチェックできる。

見える化の効果
ルールとの適合性のチェックは、品質保証・説明責任に直結する

脆弱性・堅牢性、バイアス・公平性対策の面
重要なルールに最低限従っていることで業務適用時の影響をコントロールできる

AEI

情報をN4に変換して
見える化したもの



ルールの例

- 法廷のルール（憲法、法令など）
 - セキュリティ、プライバシー、著作権、その他各種法律
- 民間で普及しているルール（常識・業界標準・規約など）
 - 世間一般
 - 倫理観、一般慣習
 - 業界一般
 - 業界基準・ガイドライン
- 現場のルール（社内や組織でのローカルルールなど）
 - 会社のルール・基準
 - 組織のルール・基準
- 社会生活の中で意識すべきルール（感情・心理など）
 - 感情面への配慮
- その他の一般的に成り立つもの（自然法則など）

AEIは「見える化」により「認証&アクセス制限」を実現する

見える化したからこそ、認証&アクセス制限をチェックできる。

コンテンツ自体の見える化 & 認証

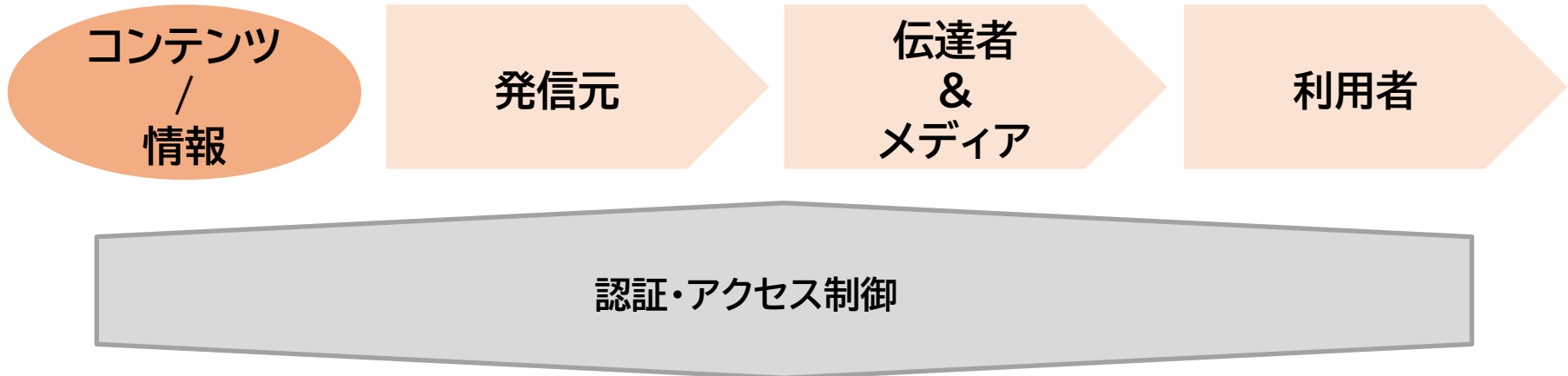
コンテンツの流通経路の見える化 & 認証

コンテンツの発信元 → コンテンツの伝達者 & メディア → コンテンツの利用者

コンテンツ、流通経路、利用者を認証した上で、

何に対して誰がどのレベルでアクセスできるかをコントロールすることができる。

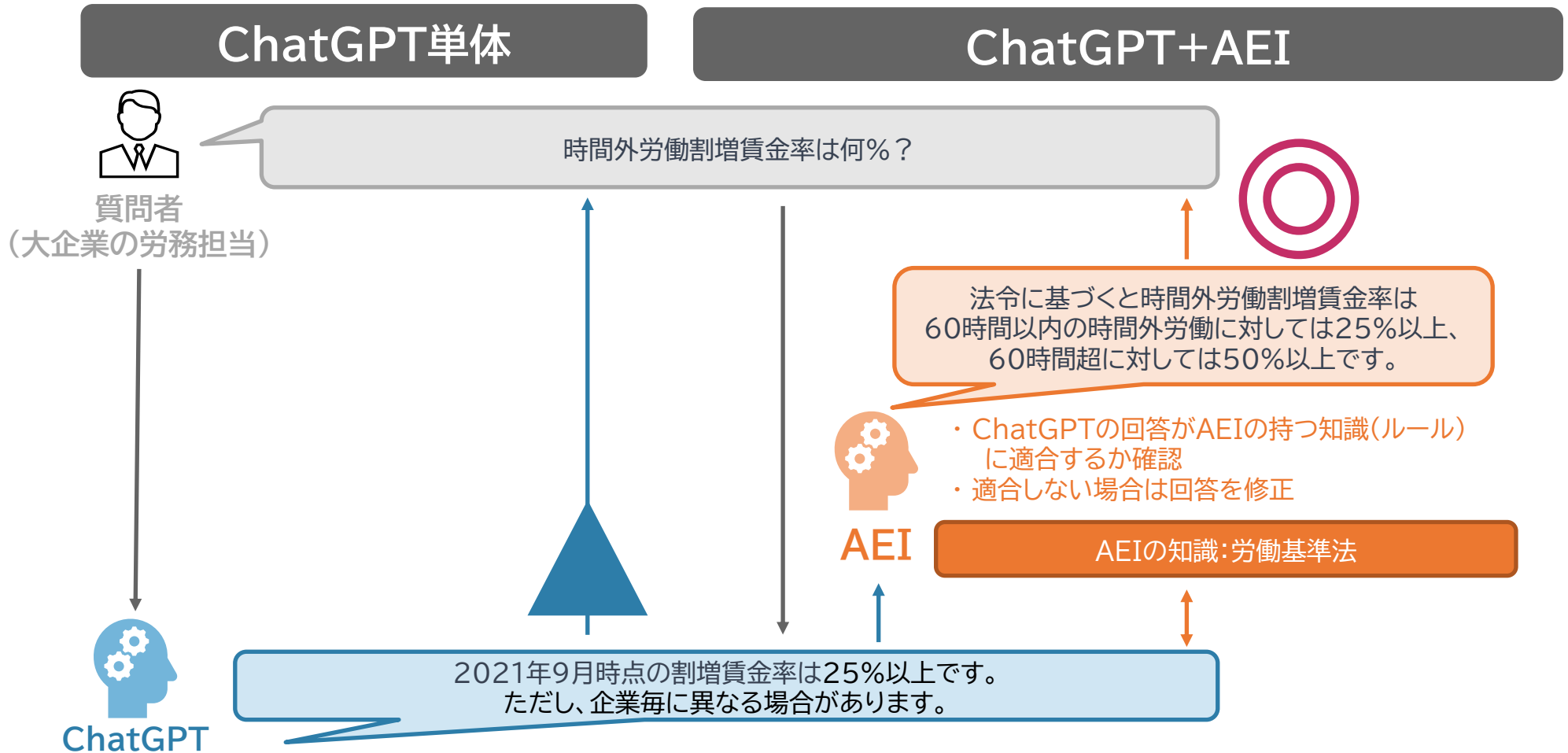
プライバシー、セキュリティー、著作権の問題は、認証&アクセス制限の仕組みがないと解消できない



AEI

AEIが法律とのルール適合性チェックを行う例

- AEIがChatGPTの解答と法律とのルール適合性チェックを行い、解答を修正。



※: 2023年4月11日時点でChatGPT Mar 23 Ver(GPT4搭載)に対して質問した際の解答を抜粋したものです。

AEIが法律だけでなく規程を含めたルール適合性チェックを行う例

- AEIは言葉や知識をグラフ構造に見える化することでルールの適合性チェックを実施

法令に基づく時間外労働割増賃金率は60時間以内の時間外労働に対しては25%以上、60時間超に対しては50%以上です。

ですが、当社の場合は、就業規則により、時間外労働割増賃金率は60時間以内の時間外労働に対しては30%、60時間超に対しては60%です。



AEI

AEIの知識
労働基準法・就業規則

	労働基準法	就業規則
60h超	50%以上	60%
60h以内	25%以上	30%



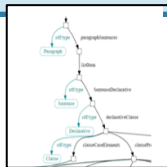
- AEIは労働基準法・就業規則の情報を予めグラフ構造に変換の上保持

- グラフ構造のChatGPTの回答とAEIの知識を比較
- 法令・規程を基にChatGPTの解答が誤っていると判断、修正



ChatGPT

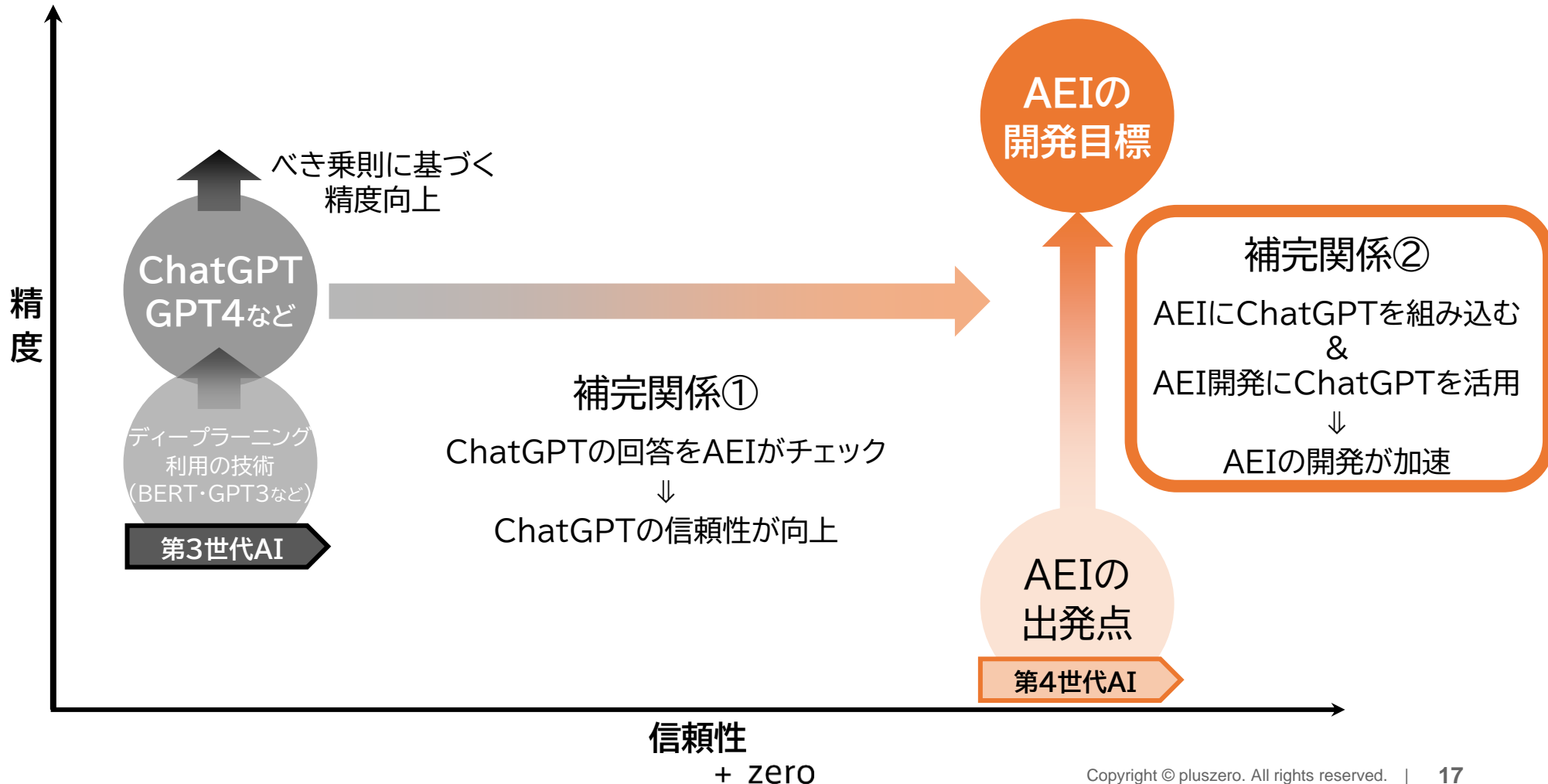
2021年9月時点の割増賃金率は25%以上です。ただし、企業毎に異なる場合があります。



- ChatGPTの回答をグラフ構造に変換

ChatGPTはAEIの開発を加速させる

- ChatGPTはAEIに組込可能
- AEI開発においてChatGPTを活用
- ChatGPTはAEIの開発を加速させることができる



ChatGPTはAEIの開発を加速させる

AEIはChatGPTを二つの観点で活用して、AEIの実現確率・精度を向上させる。

①対話のエンジンとしての利用
13の発話ロジックの一つとして活用
[2023年4月時点]

ディープラーニング
利用の技術
(ChatGPT、
GPT4など)

発話ロジックの一つ
として活用

13の発話ロジック

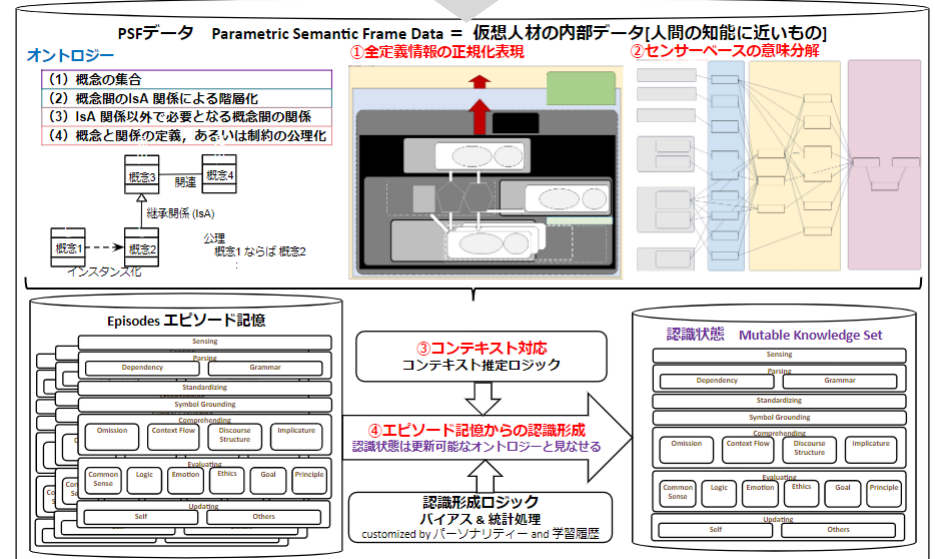
AEI
Artificial Elastic
Intelligence

②PSFデータ整備の効率向上

知能を見える化したPSFデータのデータ整備のために活用する。各サービス・モデルの利用規約等に即して大規模言語モデル全般の活用を前提。

ディープラーニング利用の技術
(ChatGPT、GPT4など)

PSFデータの整備に活用



AEIは運用すればするほど、効率的に賢くなれる

見える化の効能

見える化された情報に基づいて

局所的に動的更新できる

明示的に推論できる

- 動的更新

- 知識の獲得・更新

- 事物・人物・抽象概念

- それらの状態

- それらが関わるイベント

- ルールの獲得・更新

- OK・NG行動

- OK・NG状態

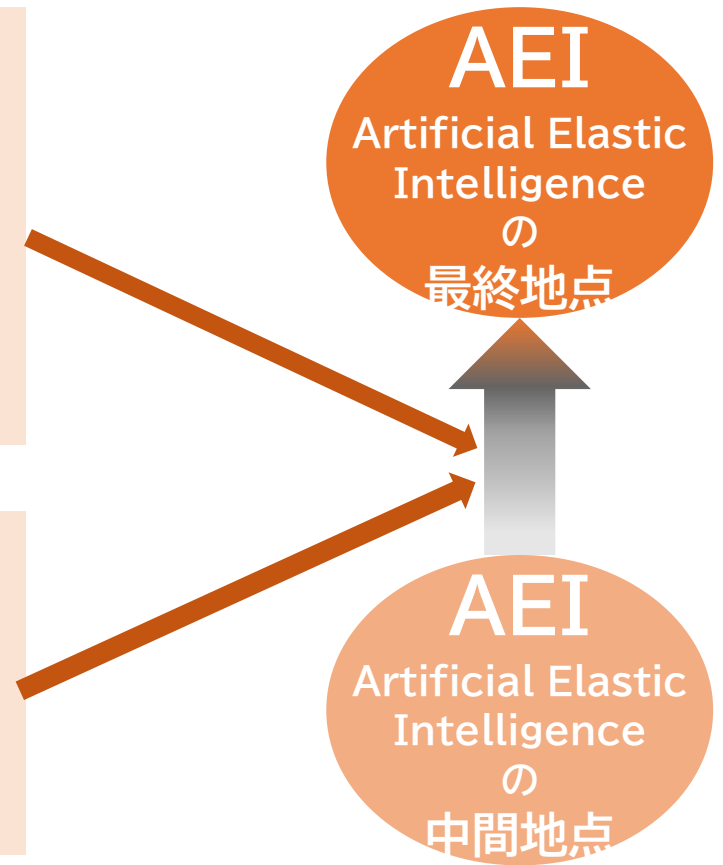
- 各種推論

- 抽象⇔具体

- 帰納・演繹・仮説推論

- 決定度合い

- 決定論的・確率論的・統計論的



AEIとChatGPTの比較

		ChatGPT	AEI	AEI+ChatGPT
AIの対応範囲		広範		業界特化 *業界拡大は26年10月期以降加速
AIの性能	精度	○	△→◎ *23/10期～26/10期に向上	○→◎ *23/10期～26/10期に向上
	信頼性	△→? *ディープラーニングの外部機能次第		◎
	現場適用力	△→? *現場依存・プラグインの進化次第		◎
AIの開発難易度	費用面	△	○	◎ *知能の整備効率向上
	人材面	理系人材中心 *高等数学、計算機科学など		文理融合人材中心
	技術面	大規模言語モデル	意味の見える化	意味の見える化 +大規模言語モデル

参考文献

- *1: Bloomberg. “Microsoft Invests \$10 Billion in ChatGPT Maker OpenAI”.
<https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai?leadSource=uverify%20wall>, (参照 2023-04-11)
- *2: 日本経済新聞. 「学校のChatGPT指針、「文科省が策定」 松野官房長官」.
https://www.nikkei.com/article/DGXZQOUA062FS0W3A400C2000000/?n_cid=SNSTW006&n_tw=1680752661, (参照 2023-04-11)
- *3: OpenAI. “GPT-4”.
<https://openai.com/research/gpt-4>, (参照 2023-04-11)
- *4: Wei, J., Wang, X., Schuurmans, D., Bosma, M., Chi, E., Le, Q., & Zhou, D. (2022). Chain of thought prompting elicits reasoning in large language models. arXiv preprint arXiv:2201.11903.
- *5: OpenAI. “GPT-4”.
<https://openai.com/research/gpt-4>, (参照 2023-04-11)
- *6: Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. Advances in Neural Information Processing Systems, 35, 27730-27744.
- *7: OpenAI. “GPT-4 Technical Report”
<https://arxiv.org/abs/2303.08774>, (参照 2023-04-11)
- *8: OpenAI. “New and improved content moderation tooling”
<https://openai.com/blog/new-and-improved-content-moderation-tooling>, (参照 2023-04-11)
- *9: OpenAI. “GPT-4 System Card”.
<https://cdn.openai.com/papers/gpt-4-system-card.pdf>, (参照 2023-04-11)
- *10: OpenAI. “Terms of use”.
<https://openai.com/policies/terms-of-use>, (参照 2023-04-11)