情報学の未来の問いを総合知によりデザインする取り組み

宇野 毅明　（国立情報学研究所
＆総合研究大学院大学）

http://research.nii.ac.jp/~uno/index-j.html
e-mail: uno@nii.ac.jp

2023年 5月25日 木曜会合・総合知活用事例

国立情報学研究所　情報学プリンシプル研究系 教授

専門分野

アルゴリズム理論 と データマイニング

情報処理学会、電子情報通信学会、OR学会

ファンド：代表・分担

さきがけ（AI）、CREST（触媒）、基盤A（SNS）

学術変革A（情報）、Q-LEAP（量子計算）

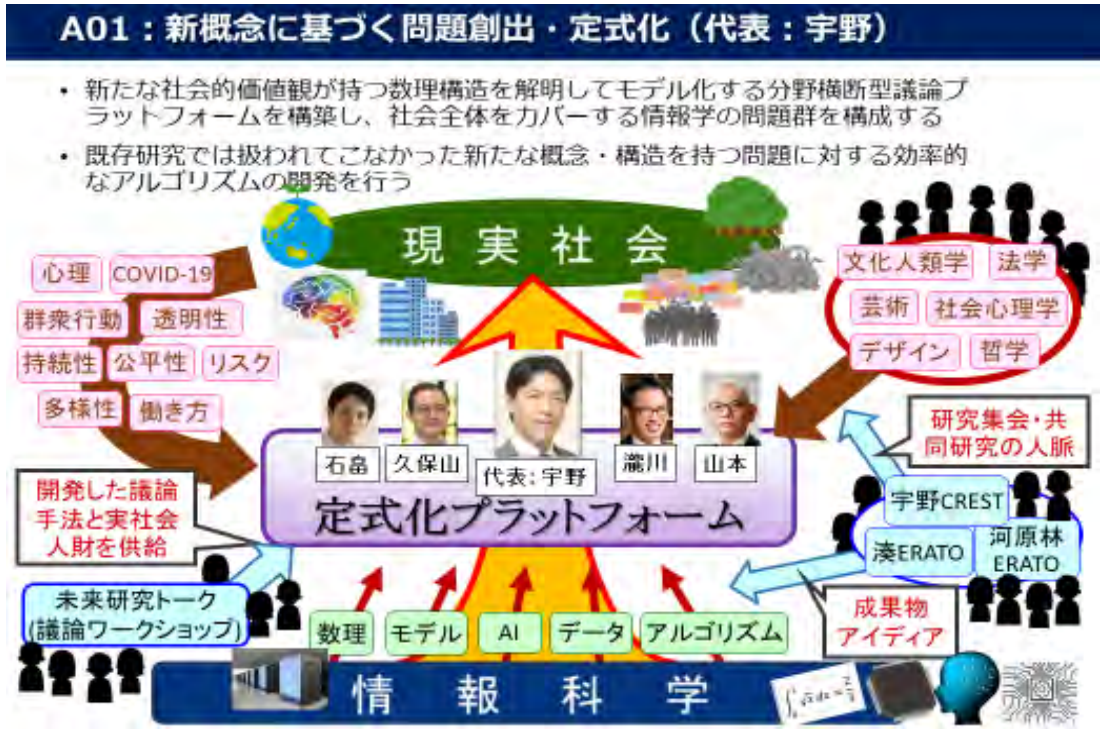
ムーンショット（目標3）、さきがけAD（信頼できるAI）

共同研究など

物作り、工程管理、物流、労務管理、コールセンター

婚活、ファイナンス、広告、腸内細菌、コンサル、…

学術変革A アルゴリズム基盤 (http://afsa.jp)



＋ 今後、情報学が目指すべき、新しい問いの群を整備し定式化する
（問題カタログ作り）

＋ 問いを多様な視点から多角的に議論し、議論手法、議論のプラットフォームを構築する

＋ 上記の問いの群を他班と共同で研究する

情報学の新しい問いとは

＋ 過去から考えられてきた古典的な情報科学の問い
＋ 最近考えられている流行の問い
＋ 将来的に起こりうる新しい領域の問い

問い

＋ ユーザサイドから見た要求、利便性、信頼性など
＋ 現代社会の新しい概念、公平性、マイノリティなど
＋ 社会心理など、環境と社会に大事で既知の知見の導入
＋ SNS、デマ、パニックなど、新しい現象の考慮・対策

概念

問題意識： デマ情報の検知と周知では、トイレットペーパー買い占めは防げない

例） 機械学習は精度を追求。倫理を守るには倫理を制約や精度に落とし込む（手法自体はまったく変わらない）

→ 新しい倫理の概念・世界観を生む手伝いをする機械学習は？

新しい問い作りのために

- このような新しい問い立ては、情報学の研究者だけではできないだろう

  特に、人文系の研究者、産業の人々との議論が必要

- しかし、人を集めただけでは、自分の立場から、自分の知識を話し、相手の情報を得るだけになる

  自分たちの専門の外側で、自分たちのディシプリンにないことを、考え、発想することは難しい

融合研究の様相

- 単に融合研究と言っても、いろいろな形がある
  大きく4つに分かれそうである

1：道具として利用

OCR、DB検索、数学定理の利用、人文の知見の参照

2：メソッド・作業の輸入

人文DB作成、DS型自然科学
クラウドソース、ビッグデータ

3：トピックの輸入

自動運転の倫理
いじめ問題のツイート解析

4：根源的な方法論、あるいは学問分野の基盤となる（暗黙の）前提の破棄と構築

プロジェクトの進め方

1. 理系・文系の研究者（両分野から4人ずつ）を集め、宇野の実践から得ている「良い議論のための仮説」を実践して議論を行う

2. 議論や交流で行われていることを観察、分析し、効果的な文理融合の手法化する

3. 新しい問いのための概念構築、具体的なモデル化と定式化を行う

4. 取り組める問題について（共同）研究を行う

交流・議論手法の構築

- 研究内容の紹介 → 興味、熱意、動機など内面へ

- きれいに整った議論するための場・空間
→ 多少雑多で、自分の所有感があり、何をしても良い、へ

- 目標は成果 → 議論すること、仲良くなること、考えること、見出すこと、そして、自身が成長することへ

- 自身の専門性を活かす → 思考、伝達、発想の力を活かす、へ

- 遡る問いインタビューによる、研究構想の可視化・構造化
- 興味と熱意の大きな偏りに対応した議論形式設計法
- 議論場の設計手法（ハード、ソフト）
- 効果的な議論の制御方法

ファンドセミナー、さきがけ研究者交流、婚活イベントに応用

ソーシャルメディア上の誹謗中傷の研究

「ひどい誹謗中傷だ」「騒ぎすぎである」など個々人の肌感覚に沿った漠然とした印象論に終始し、人それぞれ認識が違う

- 全体を俯瞰しメカニズムを知ることが重要： 何が“誹謗中傷”と呼ばれ，どのくらいあり，どのように表現されているか

  → より正確な像を提示し、対策へとつなげる

表現と書き手の意図や言葉の効果の読解に長けた文学の研究者と、概念を整理しての構造化に長けた哲学の研究者が情報学的アプローチで分析する

データ機能的な類型構築（クラスタリング）のやり方にそって、似た表現や意図のコメントを集めてカテゴリを作り、各カテゴリの数を数えて全体を俯瞰し、関係性を分析し、新たな仮説を得る

何が新しいか

哲学から見ると

+ 既存の哲学的概念を現実の事例に当てはめる個別研究が中心

文学から見ると

+ 文学作品と関連資料以外のものを研究対象にしない
+ 複雑な文章を扱うので、ラベリングのような簡単化は嫌う
+ 読解するのはあくまで作品と作家の意図であり、一般の人や平凡な群衆が書きがちなものにあまり興味はない

情報学から見ると

+ テキストの表面的な意味は扱うが、文脈は扱えない
+ 書き手の意図や心理は考えない

そして、 どの分野の研究になるのか、よくわからない

構築したカテゴリ

表現の3つのタイプとサブカテゴリ

Negative narratives

N1(a-d). 性暴力があったことは認めるものの被害者の責任を問うもの
N2(a-c). 性暴力自体があったかどうか自体を疑問視するもの
N3(a-d). 問題の周縁から，間接的に否定を試みるもの
N4(a-c). 人格や属性に対して否定を行うもの
N5. 大きな主語に向かうミソジニー

Supportive narratives

S1. 加害者を非難するもの
S2. 被害者を直接的に擁護，応援するもの
S3. 自分の経験をもとに被害者擁護を表明するもの
S4. コメントにあらわれた否定的な言説に対抗するもの

Other narratives

O1. 事実をただ述べるもの
O2: 自分や知り合いの経験の話をするもの
O3: 政治談義や思想信条など大きなものに結びつけて語るもの
O4: 対象の内側や裏側を推察して自分の見立てや考えを述べるもの
O5: 類似する事柄やその構造に言及するもの
O6: その他，SでもNでもO1-5のいずれにも入れられないもの

コメントを精読して
表現の形式と否定の根拠
に基づいて分類を構築

カテゴリとコメントの例

Label	N: 否定する表現 Negative narrative
	N1(a-d): 性暴力があったことは認めるものの,被害者の責任を問うもの
N1a	なにかしらの利得を得ていたとするもの *「役に実力が見合ってないのに、体使って役を得たケースもあるはず。それを性暴力っていうのはしっくりこない」*
N1b	なんらかの不注意・過失があったとするもの *「薄々、わかっていたことだと思います。舞妓から芸妓に襟変えするまでに嫌なら逃げ出したらよかったのではないですか。影があったりして尚華やかな芸妓の世界になって行くのだと思います。芸能界、アイドル、女優になるにも肝が座ってないとなれないと思います」*
N1c	すべての人が被害を受けているわけではないとするもの *「1人の元舞妓さんの言葉は本当だとしても、業界全体がそうだと直結させるのは早計だと思います。そして、それを煽っているのはメディアでしょう？」*
N1d	告発時期の妥当性を疑問視するもの *「こういう告発は、結局有名になった人が後から声をあげるから、そんなこと言っても恩恵は受けたんじゃないの？と思われてしまう。多くの人が関わった作品がどんなにぶち壊しになろうと、その場で騒ぎ立てるのが一番良いんじゃないかと思う。加害者は最悪だが、これは被害者も自分の意思である程度コントロールできるものだったと思う。女優も、無名だろうと有名だろうと、そんなことをして仕事を得ることは、結局誰のためにもならないことをよく考えるべき」*

分析と仮説構築

類型に基づいてコメントにアノテーションをする（12articles, 6 topics, 727 posts）
1) 著名人の#metoo，2) 映画界の #metoo，3) 性暴力の訴訟，
4) 舞妓の#metoo，5) 子どもに対する性暴力，6) 政治界における#metoo

	N1a	N1b	N1c	N1d	N2a	N2b	N2c	N3a	N3b	N3c	N3d	N4a	N4b	N4c	N5
Celebrity	0	0.02	0	0	0.01	0.049	0.039	0.114	0	0.111	0.026	0.178	0.143	0.254	0.057
Film industry	0.234	0.079	0.028	0.053	0	0.048	0.05	0.022	0.048	0.24	0	0.106	0	0.07	0.022
Lawsuit	0	0.1	0	0	0	0.013	0.25	0.025	0.313	0.2	0.013	0	0.013	0.025	0.05
Maiko	0.021	0.229	0.125	0.083	0.063	0.104	0.083	0.021	0	0.104	0.083	0	0	0.042	0.042
Children	0.025	0	0	0	0	0	0.5	0.025	0.075	0.175	0.125	0	0	0.05	0.025
Politics	0.048	0.095	0	0.095	0	0	0.071	0.048	0	0.214	0.119	0.024	0.071	0.143	0.071

	S1	S2	S3	S4	O1	O2	O3	O4	O5	O6
Celebrity	0.122	0.157	0.028	0.693	0.197	0.018	0.07	0.615	0	0.1
Film industry	0.369	0.144	0.042	0.444	0.049	0.018	0.08	0.49	0.23	0.133
Lawsuit	0	0.094	0	0.406	0.271	0.021	0	0.375	0	0.333
Maiko	0.364	0.273	0.091	0.273	0.018	0.018	0.036	0.764	0.073	0.091
Children	0.311	0.05	0.264	0.375	0.042	0.189	0.029	0.598	0.042	0.1
Politics	0.357	0	0.143	0.5	0	0.025	0.272	0.469	0.099	0.136

哲学と社会学の先行研究と、データが合致することを証明

＋ 著名人には人格や属性への攻撃が多い

＋ 映画界では「女優の自業自得」とされやすい

＋ 被害者が未成年だと攻撃されないが，職業（舞妓）を持っていると成人女性と同じように攻撃される

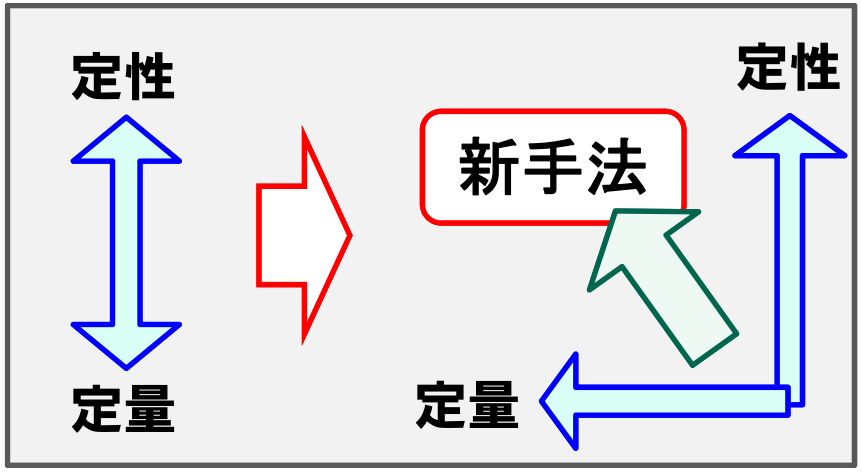
既存研究には存在しない仮説の導出

- 質的な側面を担保し、量的な概念を取り入れた人文研究の発展
- 人文的質が入ったデータの構築と機械学習での利用の促進

新しい価値の創出、苦労したこと

価値 人文学と情報学、両者の価値観と技術が混ざった手法の開発

定量⇔定性の両極化からの脱却



異分野の研究者が、興味、価値観、技術の長所を活かして研究

メタ研究手法の研究

苦労・困難

- 成果の出しどころがない
- 破棄すべき常識の発見と打破

✗ おしゃれで整った空間
✗ 参加者全員が活発な議論
✗ 分野の最先端の議論

大変 試行錯誤の連続、評価できる人がいない、PDのキャリア…

しんどい ゴールが見えない、人の雇用、リーダーシップ…

以下、参考資料

誹謗中傷の研究

#metoo に対する”誹謗中傷”的語りの類型化とアノテーション

武富 有香（国立情報学研究所 情報学プリンシプル研究系）
松田 智裕（立命館大学 衣笠総合研究機構）
須田 永遠（国立情報学研究所 情報学プリンシプル研究系）
宇野 毅明（国立情報学研究所 情報学プリンシプル研究系）

概要： #metoo 運動以降，ソーシャルメディア上にあらわれるようになった誹謗中傷のメカニズムとオンライン上の言説空間の全体像を理解するために，性暴力のトピックについて語る人々の書きぶりや言い回しにそって語りを類型化し，コメントにラベルをつけて分析を試みる．まず，性暴力の告白やその暴力の被害者に関する Yahoo!Japan ニュース上の記事に対するコメントを読み，類型を作成する．この類型にしたがってコメントにアノテーションを行い結果を観察すると，得られた結果と社会学および哲学の既存研究の知見との対応が見出せると同時に，その既存研究に新たな仮説を付け加えることができる．本研究では質的に読まれていたコーパスを量的に扱えるようにする情報学的なアプローチをとる一方で，類型化とアノテーションの作業，すなわち精緻にテクストを読む技術が必要な部分は人手で行う．この手法を用いることで類型に質的な深みを与えることができ，この明確な解釈性を持った類型によってコーパスを”測る”ことが可能となる．

キーワード：ナラティブ，ミソジニー， #metoo，アノテーション，ソーシャルメディア，誹謗中傷

Categorization and annotation of misogynic narrative against #metoo

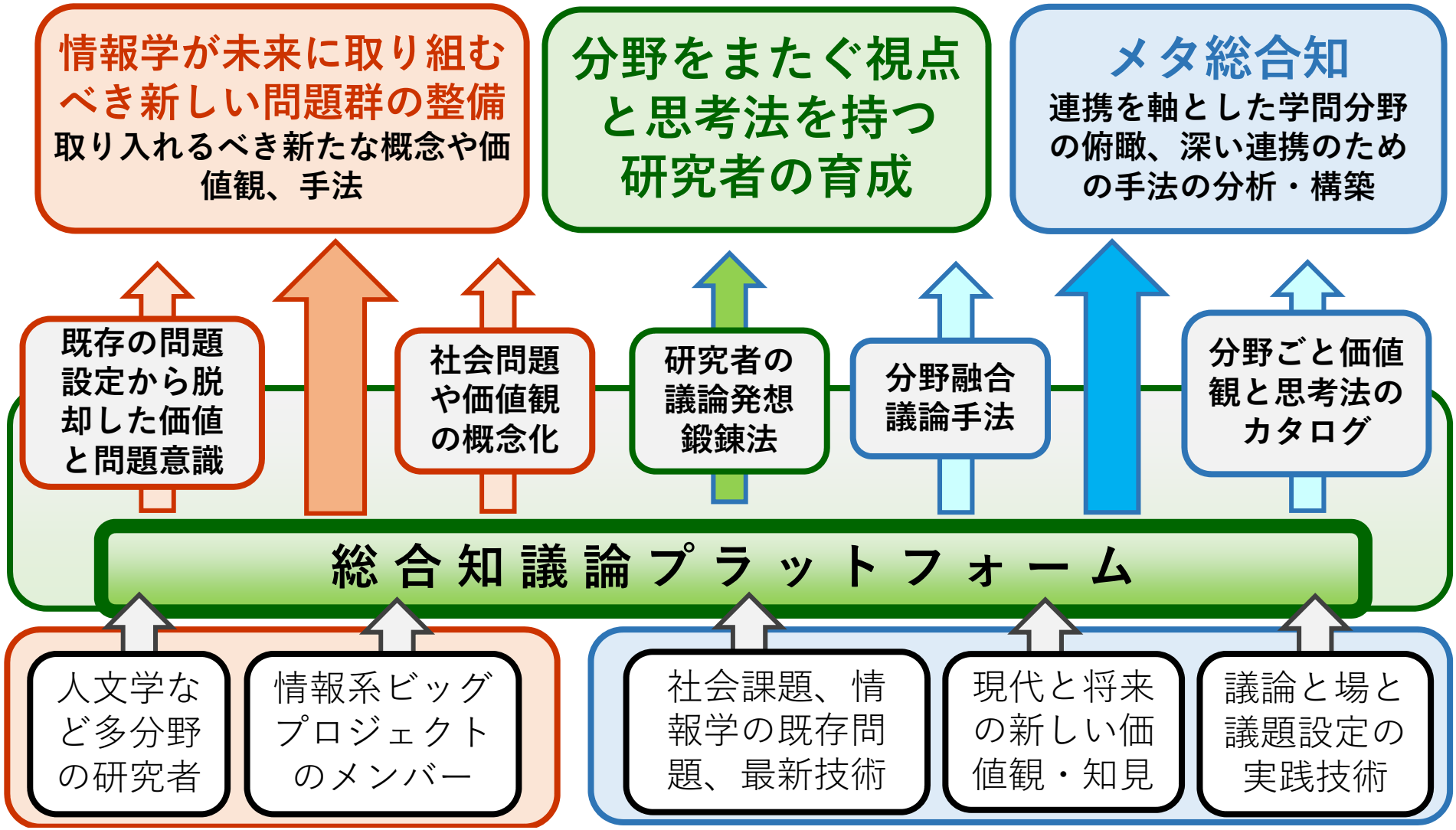
Yuka Takedomi (National Institute of Informatics)
Tomohiro Matsuda (Ritsumeikan University)
Towa Suda (National Institute of Informatics)
Takeaki Uno (National Institute of Informatics)

Abstract: We created a typology of narratives about sexual violence in social media that have emerged since the #metoo movement, according to the writing styles and phrases, and labeled the comments to understand the overall picture of the online discourse space and the mechanism of this situation. First, we categorized the comments on Yahoo! Japan News that users wrote in response to the news articles about sexual violence and victims of such violence. Then we annotated them according to the typology and analyzed the results. We not only found correspondences between the results and the previous research in sociology and philosophy but also could add new hypotheses. While this study takes an informatics approach to make a qualitatively read corpus quantitatively

情報処理学会じんもんこん2022論文集
213-220

情報学の未来の問題を総合知によりデザインする取り組み

多分野の研究者が互いの価値観や思考法を取り入れ、情報学、特にアルゴリズム理論の新しい問題群を新しい視点、新しい手法で捉え直しデザインすると共に、多分野の専門性や思考、価値観の差異を分析し、総合知的な思考の手法論を構築する



- この取り組みは、学術変革領域A「社会変革の源泉となる革新的アルゴリズム基盤の創出と体系化」の計画班として実地されている
- 社会心理やSDGsなどの新しい知見や概念に基づき、情報学の既存の問題を新しく設計し直すとともに、将来解くべき問題を新しい視点からデザインする
- 情報系と人文系を主体とする研究者が、互いの価値観、思考法を取り入れ思考し、議論することで、現実性を持ちながら新たな価値を生む課題設定を行う
- これらの議論の構造を分析することで、学問分野の価値観や思考法の差異をあぶり出し、融合型の議論を行うための手法論を確立する
- これらの議論と思考により多分野にまたがる視点と思考法を持つ研究者の育成を目指すと共に、このような総合知的な成長を得るための、鍛錬の方法を構築する。

人員と基礎技術

- ERATO(湊真一)、ERATO(河原林健一)、CREST(宇野毅明)、新学術領域(渡辺治)など情報系ビッグプロジェクトの人員が参画している
- 2011より続いている、分野横断型の研究力、議論力を鍛える議論道場、未来研究トークで培われた議論手法、議論分析手法を用いている

オリジナリティ

- 今の技術を展開して問題を作るものではない。新しい概念、視点、価値観のもとで既存の問題、将来の課題が情報学的な技術で取り組めるかどうかを考察する
- 情報的に取り組むことが難しい将来課題を、それぞれの分野の視点で記述するものではない。他分野の考え方をお互い交換し、思考法から構築するものである
- 自動運転やリーガルテックのように、産業が活発になってから総合知的な議論するのではなく、将来の産業のために先回りで議論したい。