

ファクトチェック・イニ
シアティブ・ジャパン

FIJ + SmartNews + 東北大による ファクトチェック支援技術の 実証実験構想

乾 健太郎

東北大学 / 理化学研究所

ファクトチェック・イニシアティブ・ジャパン

インターネット

2017年06月21日 12時07分

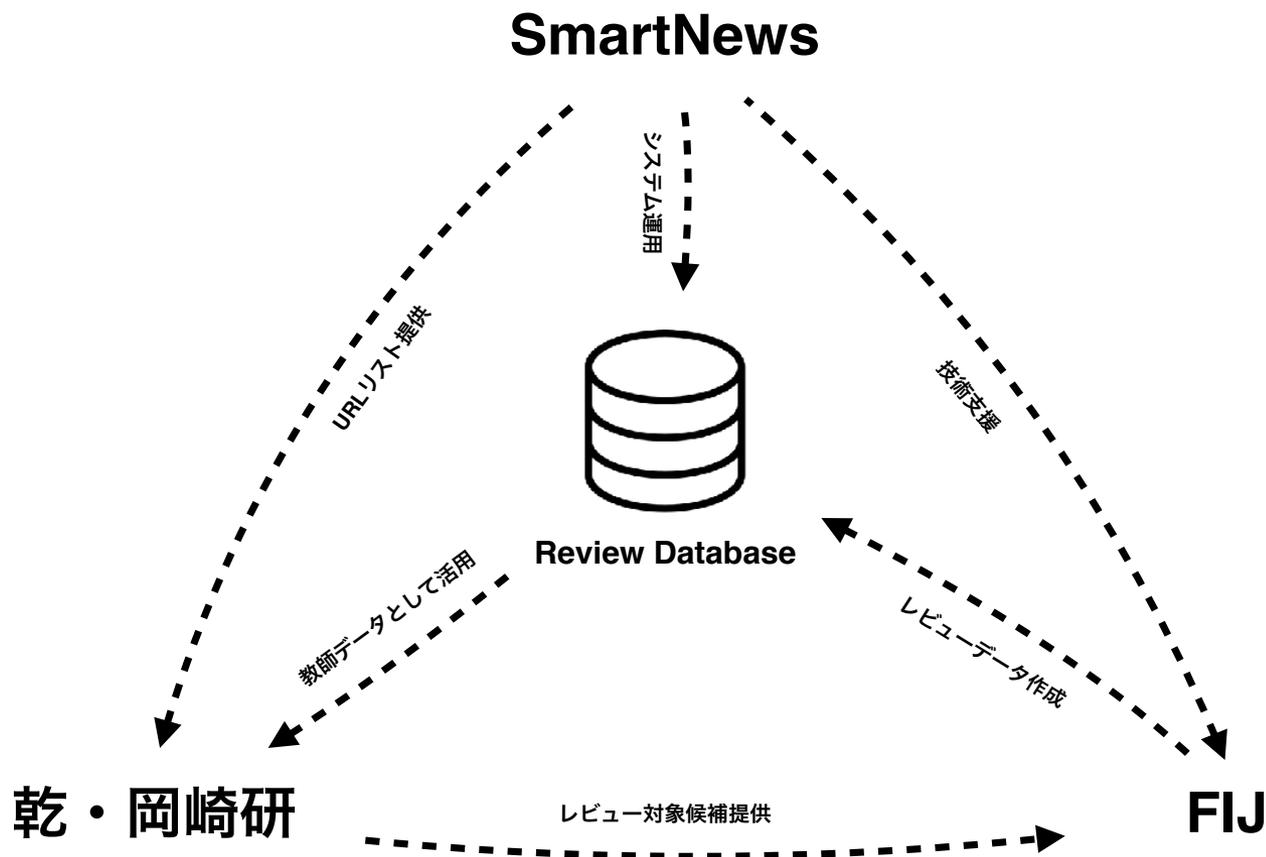


フェイクニュース対策で「ファクトチェック」団体発足、機械学習
を利用した支援目指す

https://www.bengo4.com/internet/n_6252/

ファクトチェッカーを支援する仕組みづくり

取り組み当初の体制案



乾 健太郎

専門分野：情報科学、人工知能、自然言語処理

1995年 東京工業大学 大学院情報理工学研究科 博士修了 博士(工学)

1995年 東京工業大学 大学院情報理工学研究科 助手

1998年 九州工業大学 情報工学部知能情報工学科 助教授

2002年 奈良先端科学技術大学院大学 情報科学研究科 助教授

2010年 東北大学大学院情報科学研究科 教授

2016年 理化学研究所 革新知能統合研究(AIP)センター

自然言語理解チームリーダー 兼任

ファクトチェック・イニシアティブ 理事 兼任

国立情報学研究所 客員教授 兼任

情報通信研究機構 有期研究員 兼任

言語処理学会 理事、会誌「自然言語処理」副編集長

情報処理学会 自然言語処理研究会 主査

東北大学 自然言語処理研究室

Since 2010



乾健太郎 教授
岡崎直観 准教授 (現東工大教授)
井之上直也 助教
折田奈甫 助教(兼担)
松林優一郎 特任助教
田然 特任助教
水本智哉 AIP特別研究員
Paul Reisert AIP特別研究員
松田耕史 研究員

技術・事務スタッフ： 山口健史 高橋容市 磯部順子 菅原真由美 井上幹子

博士学生：8 修士学生：15 学部学生：5 研究生：1



言語処理による情報・知識の編集

報道メディア

議論支援

法務

文献

マーケティング

コールセンター

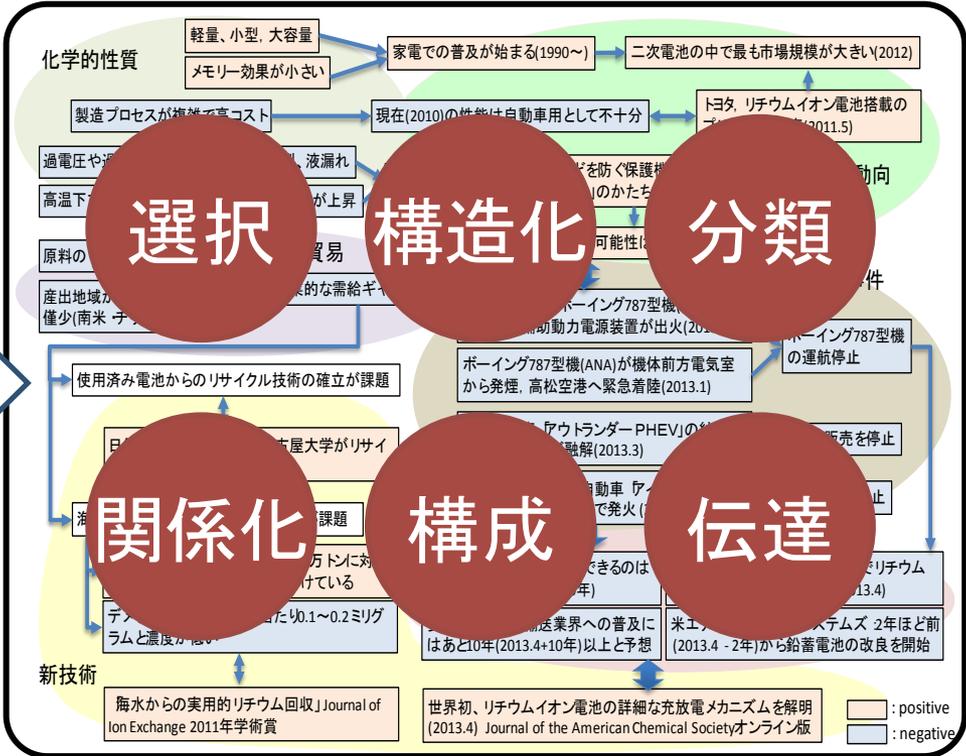
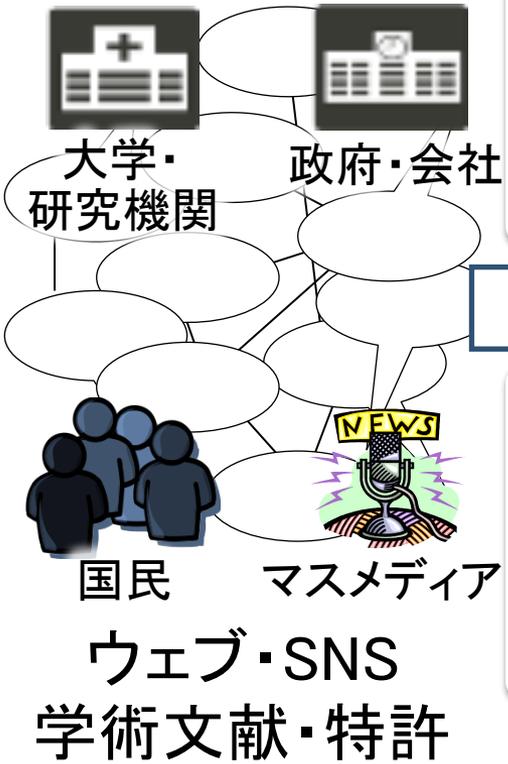
医療・ヘルスケア

↑ アプリケーション

言語解析
意味解析
リンク解析

解析・集約

情報検索
知識獲得
談話解析



情報・知識の整理・集約・可視化・分析

ネット上の誤情報の収集・集約・可視化

SNS上の様々な言明



東北太郎 @tarotohoku

【拡散希望!!】コスモ石油の爆発で有害な雨が降るらしい

誤情報候補の抽出

イソジンと被曝予防に効果がある
イソジンと被曝予防に効果がある
イソジンと被曝予防に効果がある

コスモ石油の爆発で有害な雨が降る
この震災は米国の地震攻撃だ

誤情報候補の集約

デマ情報一覧

トピック	スコア	命題
田尻智さん	18.7	海外で、ポケモンの生みの親の田尻智さんが亡くなった
コスモ石油千葉製油所	11.8	@市原市のコスモ石油千葉製油所LPGタンクの爆発により、千葉県、近隣圏に在住の方に有害物質が雨などと一緒に飛散する
女性暴行	10.4	@「阪神大震災の際には女性暴行が増えた」
有害物質	8.6	コンビニート火災に関し『有害物質が発生し、雨に混じって降ってくるので肌をさらさないように』
尾田栄一郎先生	5.3	尾田栄一郎先生が15億円を寄付
円資産	5.2	池田経済産業副大臣は、「日本企業が円資産を買い戻している
空中投下	3.0	『物資の空中投下が日本で認められていない』
サーバルーム	1.9	サーバルーム、ラックで腹部を怪我した
イソジン	1.6	被曝予防にイソジンを飲むと良い
茨城県知事	1.2	茨城県知事災害派遣要請出てない
食塩	1.0	「食塩にはヨウ素が含まれており被曝の防止に効果がある」

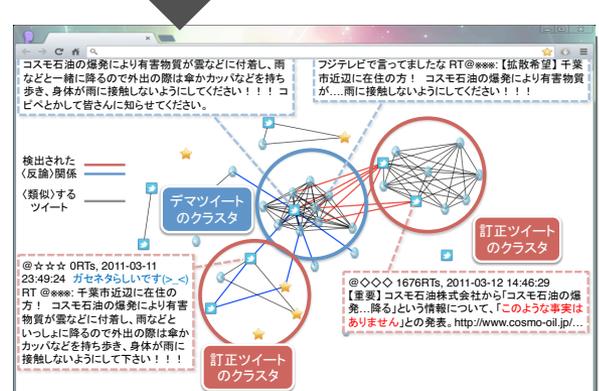
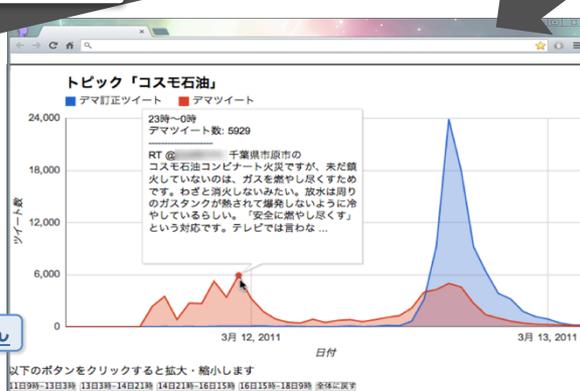
様々な視点で分析

質問: イソジンで被曝を**防**げる?

肯定: 39 (76%)
否定: 12 (24%)

被曝を**軽減**するにはイソジンがいい

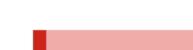
イソジンは被曝**対策**には使えません



ネット市民の大規模世論調査 (Sasaki+ ACL 2017)

Twitterユーザー100万人 × 500トピックの調査へ

対象トピック [賛成\(+予測\)](#) [反対\(+予測\)](#) 賛否が類似しているトピック

TPP		社会保障改悪/ヘリパッド建設/平和憲法改悪/盗聴法改悪/刑事訴訟法改悪/日中韓FTA/戦争法案強行
憲法改正		リビック/特定秘密保護法施行/9条改定/原発再起動/憲法9条改正/寝台特急廃止/東京オリンピック20
原発再稼働		再稼働/集団的自衛権閣議決定/原発再稼働/すべての原発再稼働/川内原発再稼働/核燃サイクル/高浜
オスプレイ		戦争する国づくり/辺野古新基地建設/沖縄新基地/伊方原発/柏崎刈羽再稼働/島根原発再稼働/日印原
性転換		一妻多夫制/刹処分/国際結婚/資金援助/移住/人材派遣/tx/大阪万博/A案/電子化
消費税増税		渡辺美樹ワタミ擁立/日韓通貨スワップ延長/人権救済機関設置法案/日韓スワップ延長/日韓通貨スワ
安保法案		安保関連法案/集団的自衛権閣議決定/特定秘密保護法案/安保法政/集団的自衛権行使容認/核燃サイ
秘密保護法		特定秘密秘密保護法/伊方原発再稼働/女川原発再稼働/全ての原発再稼働/すべての原発再稼働/日印
ヘイトスピーチ		派遣法改悪強行採決/ヘイトデモ/9月14日京都タワー下ヘイト街宣/戦争法案/差別デモ/10月5日秋葉
東京オリンピック		日米ガイドライン改定/皇室制度/違憲戦争法案/築地市場移転/殺処分機導入/核燃サイクル/生活保護
川内原発再稼働		高浜原発再稼働/全ての原発再稼働/伊方原発再稼働/派遣法改正案の採決/違憲戦争法案/特定秘密秘
ヘイトデモ		差別デモ/戦争立法/4月27日名古屋差別デモ/11月4日京都の差別デモ/9月14日京都タワー下ヘイト街
アベノミクス		東京オリンピック2020開催/寝台特急廃止/自由民主党/マイナンバー/戦争原発/特別秘密保護法/自転車の
外国人参政権		韓国人学校建設/外国人労働者受入/日韓通貨スワップ協定/反差別基本法/韓流門番服/朝鮮学校無償

ファクトチェック・イニシアティブ・ジャパン

インターネット

2017年06月21日 12時07分



フェイクニュース対策で「ファクトチェック」団体発足、機械学習
を利用した支援目指す

https://www.bengo4.com/internet/n_6252/

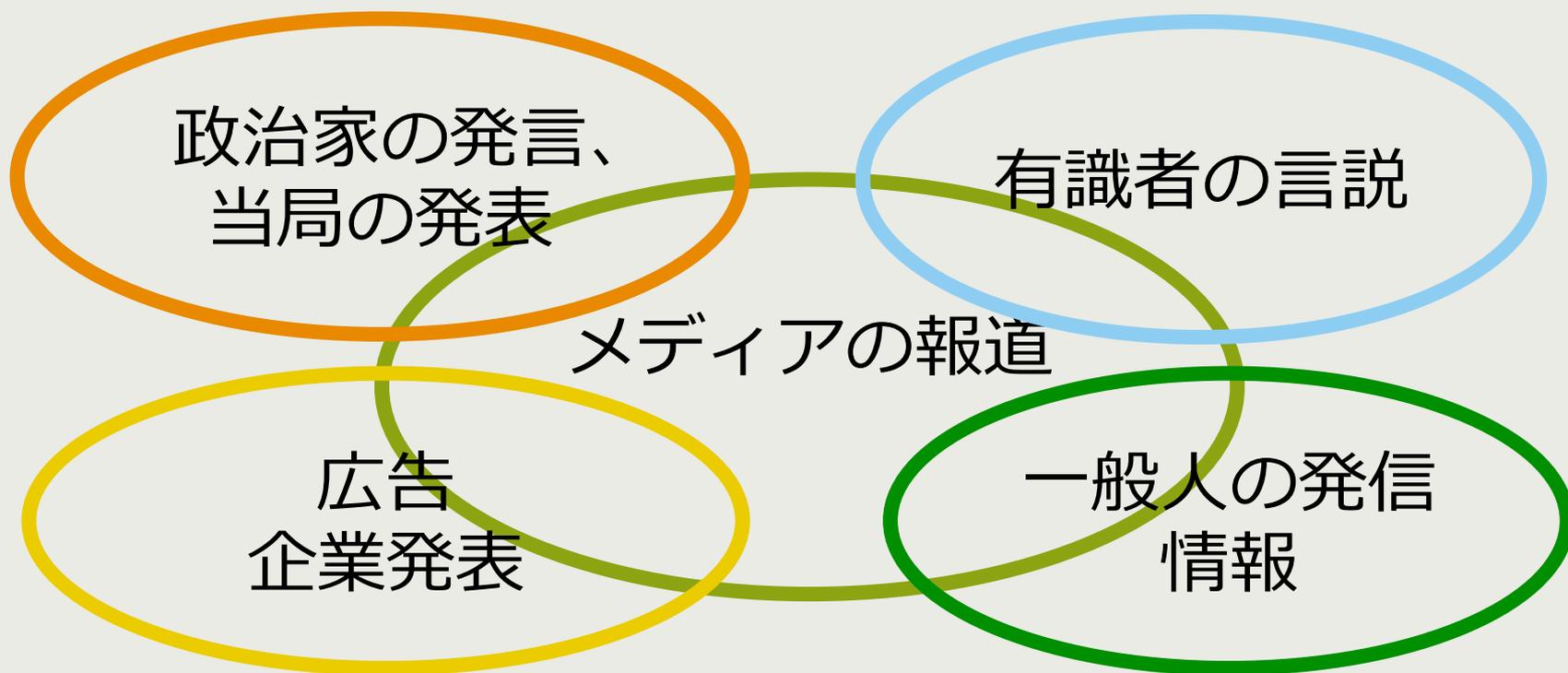
なぜ今ファクトチェックなのか？

ファクトチェックとは

- (1) 事実に関する言明の
- (2) 真偽・正確性を調査・検証し
- (3) 証拠等の判断材料を提供する

なぜ今ファクトチェックなのか？

ファクトチェックの対象



FIJを立ち上げる狙い

日本国内における
ファクトチェックの推進・普及
に向けて
業界の垣根を越えた取組み

FIJの活動方針

- **ファクトチェックの理解・手法の共有化**
 - ガイドライン・評価方法を検討
- **ファクトチェッカーを支援する仕組みづくり**
 - 支援方法を検討し、近く法人化

日本初のファクトチェックプロジェクト

FIJファクトチェック
2017総選挙プロジェクト

**Buzz
Feed**
JAPAN

 Japan In-depth

ニュースのタネ。
SEEDS OF NEWS.

GQHO

第3回FIJセミナー「総選挙ファクトチェックプロジェクトを総括する」概要説明スライドより

プロジェクトの成果(22本の記事化)

BuzzFeedNEWS

小池知事、衆院選出馬
「最初から100% ない」

不正確



産経新聞10月1日付朝刊

GoHooのファクトチェック

事実
に
反する

政令で、「前職」も
「国会議員」として
カウントされる

Japan In-depth 2017総選挙ファクトチェックプロジェクト

しんぶん読者 10月8日、2面
“法人税増えている”
虚構が真りの発見

安倍晋三氏(自民党総裁)10月8日、NHK
法人税を引き下げましたが、法人税は増えています。

どちらも「不正確」

BuzzFeedNEWS

希望の党と自民党支部
「住所が一致」は本当か?

不正確



Japan In-depth 2017総選挙ファクトチェックプロジェクト

日本選挙の専門家
松井一郎氏(大塚府知事)
立憲民主党のメンバーについて
「民進党の中心、「絶対に一言一句、動詞も含めて
露骨を責めるべきではない」という人たちの集まり」

↓
虚偽

総選挙ファクトチェック 安倍総理の解散理由に
「事実と異なる内容」



正社員になりたい人がいれば
必ず1つ以上の正社員の仕事がある

総選挙ファクトチェック 維新の「大阪で教育無償
化実現している」は「事実ではない」



維新ならできる!
増税なしで改革実現!

総選挙ファクトチェック 共産党委員長の「残業時
間の法的規制がない」は「発言に根拠は有るもの
の、その内容は必ずしも事実とは言えない」



産経ニュース9月29日

GoHooのファクトチェック

不正確

橋下氏のTwitter投稿を
無批判に引用報道。

朝日、毎日新聞は
「希望・民進の合流」を
批判的に報じていた。

BuzzFeedNEWS

立憲民主党 Twitter
フォロワー購入疑惑

根拠なし



安倍晋三首相(日本記者クラブ)

GoHooのファクトチェック

不正確

朝日新聞が
加戸守行・前安倍総理知事の
証言を取り上げたかについて
「証言した次の日に
全く(報道)して
おられない」

この発言を産経新聞議員委員
が引用し、擁護する報道

総選挙ファクトチェック 希望の党の「企業の内部
留保300兆円に課税して代替財源」は「事実と認め
るには不確かな要素がある」



日本に
希望を。

総選挙ファクトチェック 安倍総理が解散理由で語
った税収増「5兆円強」は本当か?



2%の引上げにより
5兆円強の税収となります

毎日新聞社説10月3日

GoHooのファクトチェック

不正確

政策パンフレットには
載っていないが、
政策BANKに
「女性活躍」の
項目があった

TBS NEWS2X(10月11日放送)

GoHooのファクトチェック

不正確

は元議員氏が希望の党に
「公認申請していた」
との偽情報

「リツイート(転載)
220万以上」

問題の投稿の転載は
「230以上」だった
TBSは誤りを認め、
訂正放送した

BuzzFeedNEWS

安倍首相、ほとんどの教科書で
「自衛隊から護衛と記述」

ミスリーディング



ファクトチェック記事の例

検証者 (Fact-Checker) : [Gaijin](#) (編集者)

検証記事 (Report) : [新党結成 野党は改定要件でカウントされない?](#)

掲載日 (Date) : 2017/10/9

対象文句 (Statement)

新党を作るには、原則国会議員5人以上の参加が必要だが、衆院解散に伴い旧衆院議員はカウントできない。 (広済優一・産経新聞記者)

媒体・場所 (Media/Location) : 産経新聞、[産経ニュース](#)

本文・掲載日 (Date) : 2017/10/7

判定 (Rating)

事実と反する

事実・根拠 (Fact/Basis)

政治資金規正法施行令、新党結成法施行令に、解散後の旧衆議院議員も改定要件の「国会議員」に算定するとの定めがある (附帯文に解説)。産経は範囲上で訂正した。

ネット上の端緒情報 → ファクトチェック記事へ

インターネット上の情報から 記事化までのイメージ

インターネット上の情報 (Level0) を情報収集ツールで誤報系情報を自動的にフィルターし、探索情報 (Level1) としてスクリーニング (表示) する。端緒情報作業員は探索情報を閲覧し、疑義・気づきが生じた情報を端緒情報 (Level2) として組織内で情報共有する。端緒情報のうち、影響度・深刻度の観点で調査妥当と判断したものを予備調査 (Level3) として簡易調査を行う。予備調査の中から記事化相当と認められる案件についてはさらに詳細調査となる本調査 (Level4) を行う。本調査の情報を元に、最終的に記事 (Level5) となる各レイヤーで作業分担、情報集約が行われている

インターネット上の情報

探索情報

端緒情報

予備調査

本調査

記事化

探索情報作業員

端緒情報作業員

予備調査作業員

本調査作業員

Level 0

Level 1

Level 2

Level 3

Level 4

Level 5

誤報の疑義・気づき（端緒情報）をSNSから収集

ニュース記事

〇〇〇〇、液晶TV国内生産撤退へ

〇〇新聞



記事に対するツイート

USER_1 @user_1

〇〇〇〇、液晶TV国内生産撤退へ
<http://〇〇〇〇〇〇.com/〇〇〇〇〇〇>
エコポイントなど政府のバラマキ政策に乗っかって破滅する例だね。

USER_2 @user_2

〇〇〇〇、液晶TV国内生産撤退へ
<http://〇〇〇〇〇〇.com/〇〇〇〇〇〇>
雇用が失われるわけではないのが唯一の救い。

端緒
情報

USER_3 @user_3

〇〇〇〇、液晶TV国内生産撤退へ
<http://〇〇〇〇〇〇.com/〇〇〇〇〇〇>
〇〇は撤退といってる。他紙と違いますが両方共裏付けしてるのかな？

USER_4 @user_4

〇〇〇〇、液晶TV国内生産撤退へ
<http://〇〇〇〇〇〇.com/〇〇〇〇〇〇>
これは日本製造業の落日を示す端的な一例であろう！

...

ネット上の端緒情報 → ファクトチェック記事へ

インターネット上の情報から 記事化までのイメージ

インターネット上の情報(Level0)を情報収集ツールで誤報系情報を自動的にフィルターし、探索情報(level1)としてスクリーニング(表示)する。端緒情報作業員は探索情報を閲覧し、疑義・気づきが生じた情報を端緒情報(Level2)として組織内で情報共有する。端緒情報のうち、影響度・深刻度の観点で調査妥当と判断したものを予備調査(Level3)として簡易調査を行う。予備調査の中から記事化相当と認められる案件についてはさらに詳細調査となる本調査(Level4)を行う。本調査の情報を元に、最終的に記事(Level5)となる各レイヤーで作業分担、情報集約が行われている

インターネット上の情報

探索情報

端緒情報

予備調査

本調査

記事化

探索情報作業員

端緒情報作業員

予備調査作業員

本調査作業員

Level 0

Level 1

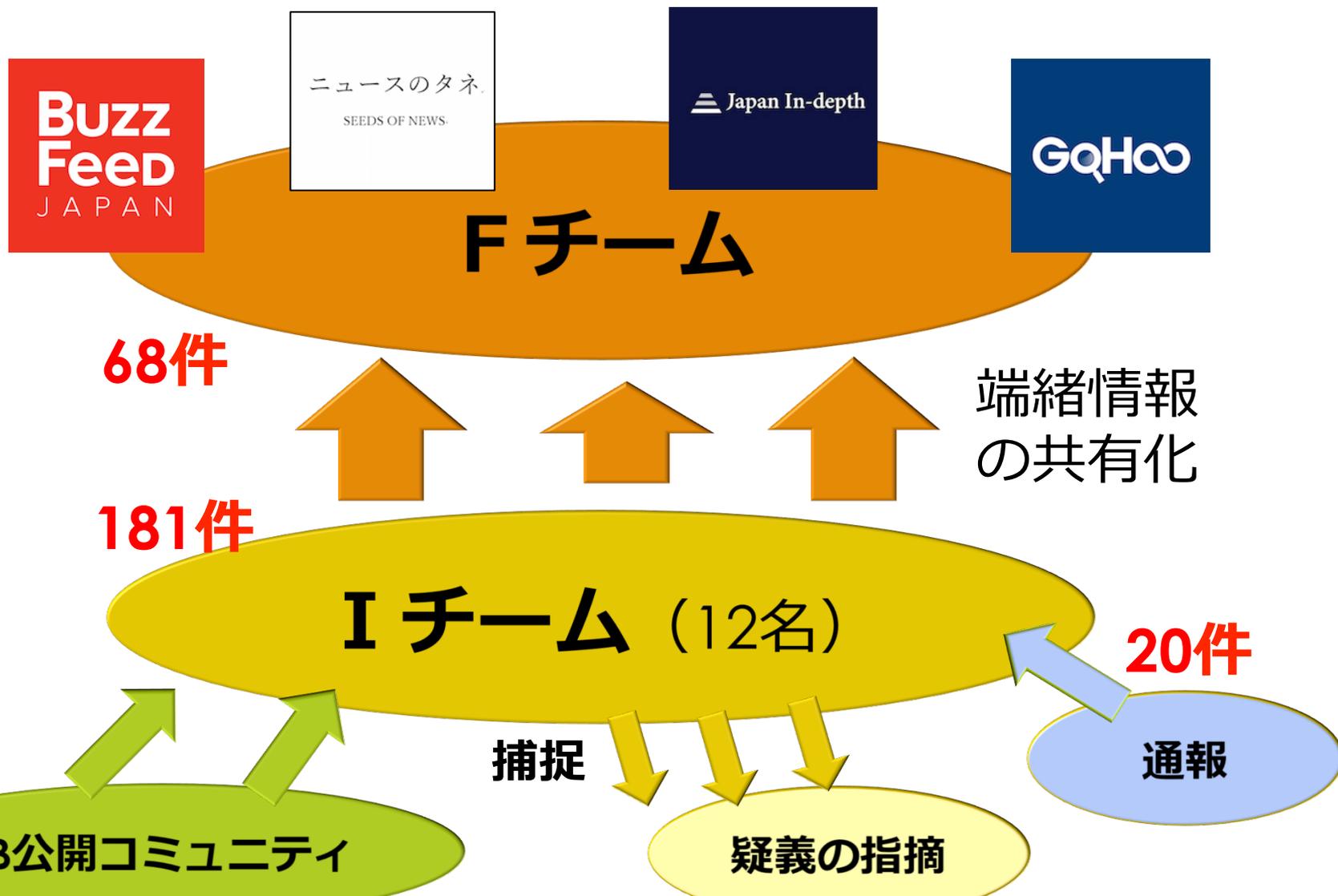
Level 2

Level 3

Level 4

Level 5

情報共有支援チーム (Iチーム)



情報共有支援チーム (Iチーム)



誤報の疑義・気づき（端緒情報）をSNSから収集

ニュース記事

〇〇〇〇、液晶TV国内生産撤退へ

〇〇新聞



記事に対するツイート

USER_1 @user_1

〇〇〇〇、液晶TV国内生産撤退へ
<http://〇〇〇〇〇〇.com/〇〇〇〇〇〇>
エコポイントなど政府のバラマキ政策に乗っかって破滅する例だね。

USER_2 @user_2

〇〇〇〇、液晶TV国内生産撤退へ
<http://〇〇〇〇〇〇.com/〇〇〇〇〇〇>
雇用が失われるわけではないのが唯一の救い。

端緒
情報

USER_3 @user_3

〇〇〇〇、液晶TV国内生産撤退へ
<http://〇〇〇〇〇〇.com/〇〇〇〇〇〇>
〇〇は撤退といってる。他紙と違いますが両方共裏付けしてるのかな？

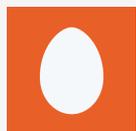
USER_4 @user_4

〇〇〇〇、液晶TV国内生産撤退へ
<http://〇〇〇〇〇〇.com/〇〇〇〇〇〇>
これは日本製造業の落日を示す端的な一例であろう！

...

ところが、 ツイートのほとんどはファクトチェックには役立たない

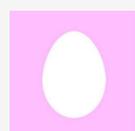
「デマ」や「誤報」で検索しても、有用な情報は1000件に1件程度



USER_1 @user_1

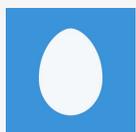
<http://○○○○○○.com/○○○○○>

○○○○でデマ言わないと呼吸出来ない
病気。



USER_6 @user_6

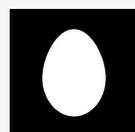
『ネットビジネスはもう稼げない！？』それは
誤報!!
こちら→ <http://○○○.com/○○○○○>



USER_2 @user_2

<http://○○○○○○.com/○○○○○>

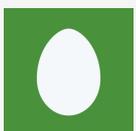
デマ吐き500円 嘘つき500円



USER_7 @user_7

<http://○○○○○○.com/○○○○○>

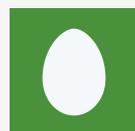
デマ吐き500円 嘘つき500円



USER_3 @user_3

<http://○○○○○○.com/○○○○○>

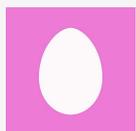
声優・○○○○、喉の病気で療養後、初
姿見せに2ch「妊娠出産説はデマ」



USER_8 @user_8

<http://○○○○○○.com/○○○○○>

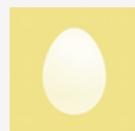
雨予報？誤報だわフハハハハハハ



USER_4 @user_4

<http://○○○○○○.com/○○○○○>

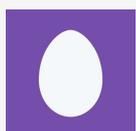
非弁行為みたいな、デマを取り締まる法
律があればいいのにな～。



USER_9 @user_9

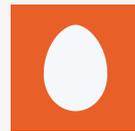
<http://○○○○○○.com/○○○○○>

オボちゃんの「オボ」は、韓国語で「誤報」
の意味。



USER_5 @user_5

★デマで金儲け★→ <http://○○○○○○.com/○○○○○>



USER_10 @user_10

<http://○○○○○○.com/○○○○○>

震度0 根室 強震モニタExtensionより。予
測ツイートの場合誤報の場合もあり

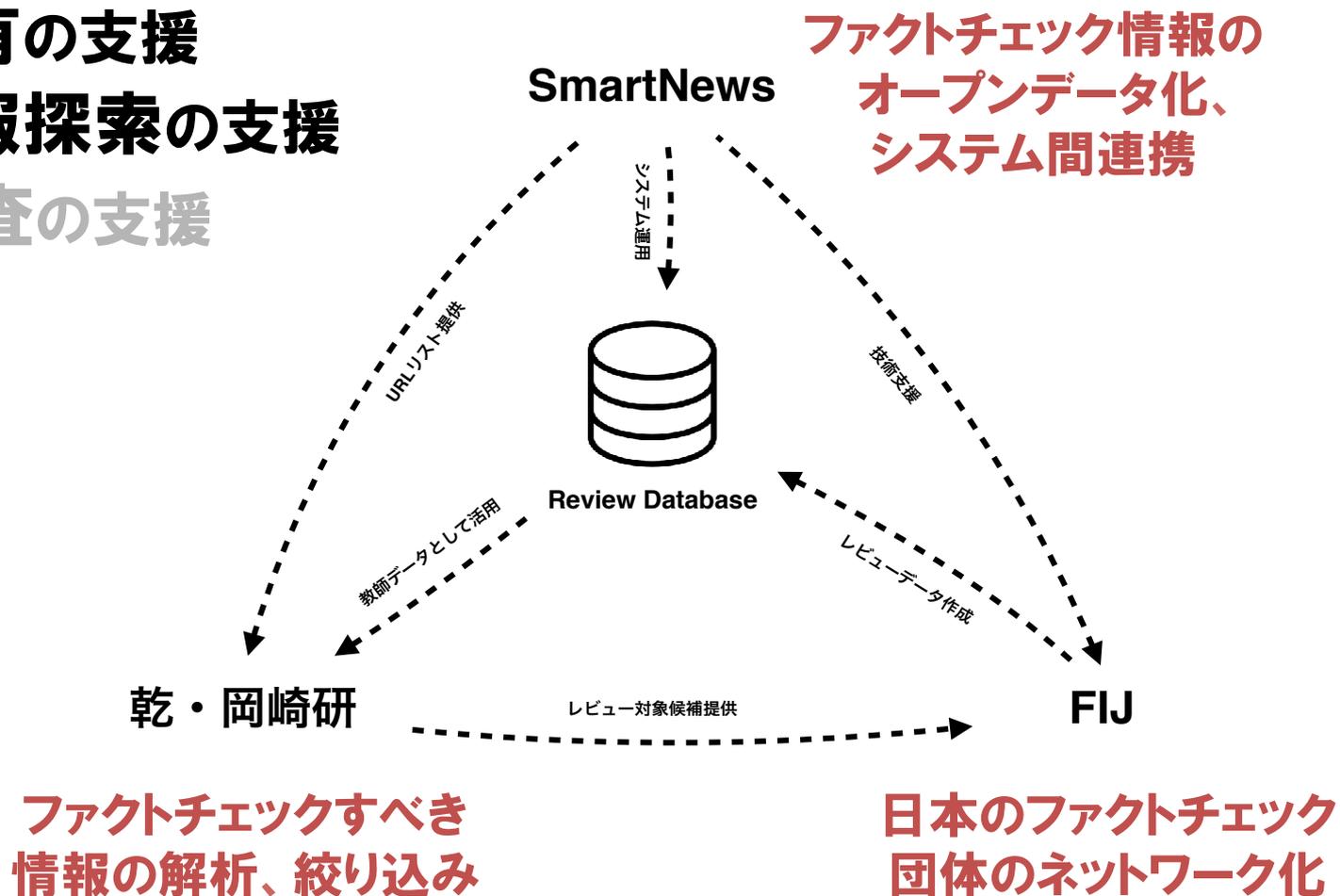
現状の作業画面の例



ファクトチェックを支援する仕組みづくり

取り組み当初の体制案

1. 情報共有の支援
2. 端緒情報探索の支援
3. FC本調査の支援



人間がチェックした方がよさそうな記事

端緒情報の候補を機械で絞り込む

ニュース記事

〇〇〇〇、液晶TV国内生産撤退へ



記事タイトル

〇〇〇〇、液晶TV国内生産撤退へ

記事URL

http://〇〇〇〇〇〇.com/〇〇〇〇〇〇

言及ツイート数

67

ツイート一覧

1. 〇〇は撤退といってる。他紙と違いますが両方共裏付けしてるのかな。
2. またまた〇〇が捏造報道。
3. また〇〇がガセ書いたのか。誤報を書いた分はその面積分の訂正を書け。
4. 嘘ニュースでした〜〇〇を信じちゃダメ

ランク
上位



最終的には
人間が判断

訃報：〇〇〇〇死去
〇〇新聞



記事タイトル

訃報：〇〇〇〇死去

記事URL

http://〇〇〇〇〇〇.com/〇〇〇〇〇〇

言及ツイート数

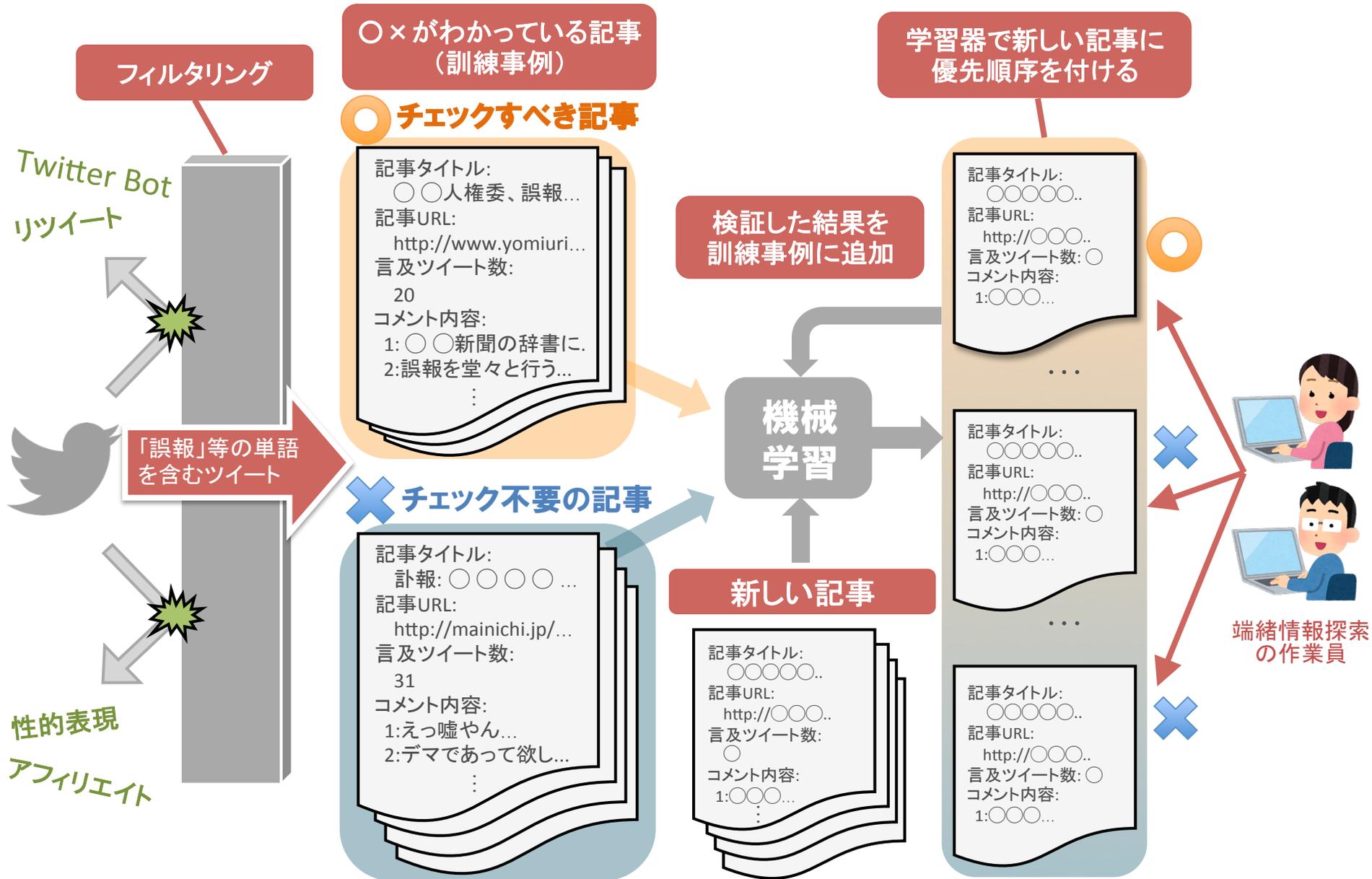
48

ツイート一覧

1. 嘘！？嘘でしょ？信じられない...(ノド)...
2. 嘘だと言ってくれよ...
3. 流れて来る情報が誤報であって欲しかった...ご冥福をお祈りします...
4. 〇〇〇〇の訃報のニュース、誤報じゃなかったか...

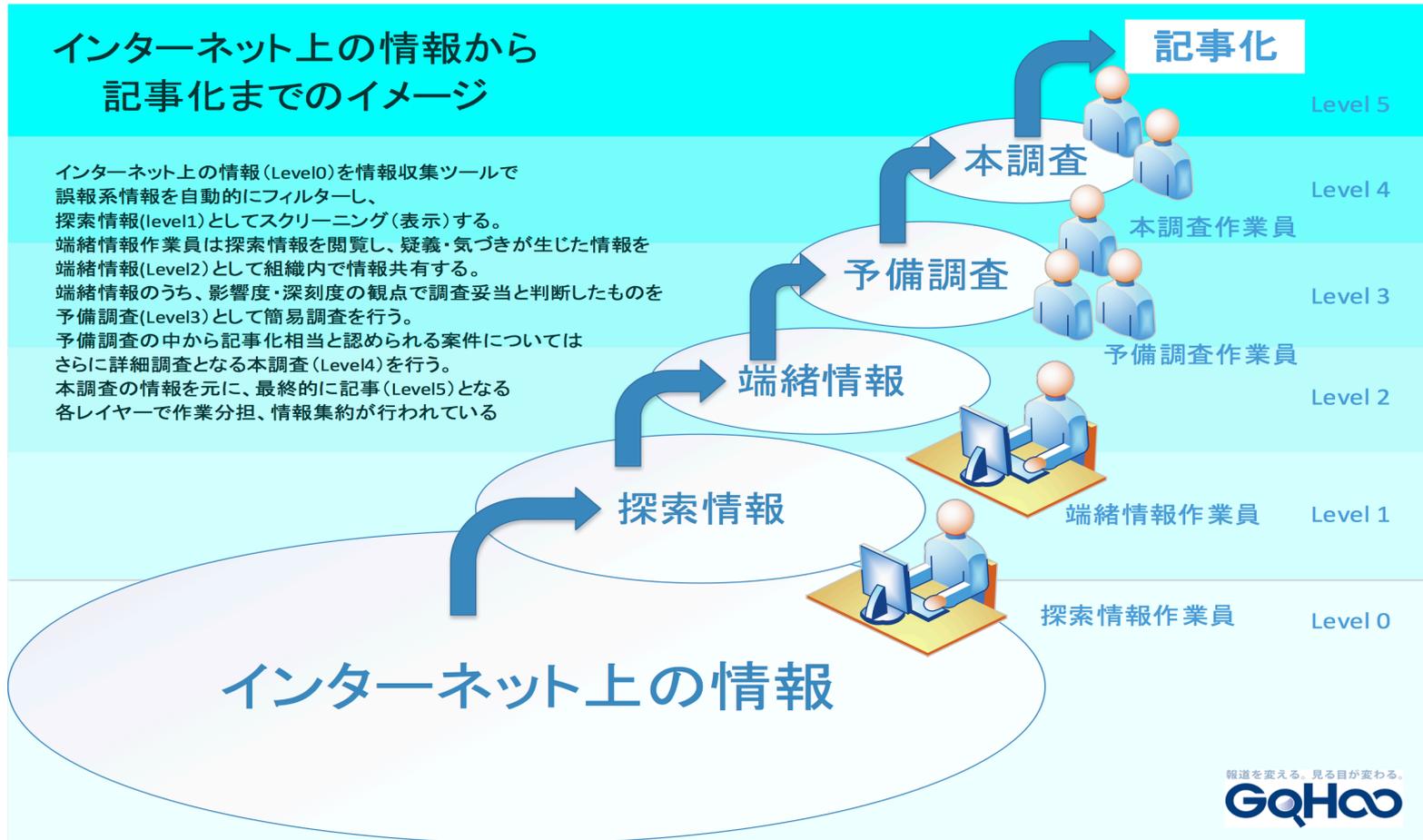
ランク
下位

自然言語処理・機械学習を使った記事の絞り込み



ねらいはファクトチェックの営みを**支援**すること

記事/言説を評価するのはあくまで**人間**



試行実験の計画

- 今年度内にプロトタイプシステムで試行実験
→ 2018年夏の実用化をめざす
- 当面の対象はニュースのみ
→ 政治家の発言などは今後の課題
- 将来はファクトチェック本調査自体の支援も

めざすのはオープンなスキームづくり

- 誰でも参照できるオープンなデータベースを整備
 - FC団体間で情報共有する仕組みを持つことが重要
 - 市民からの参加を受けとめる仕組みも
- Googleなど大手プラットフォームとも連携を推進
- 賛同する企業/研究機関/NPOを拡大

