Hybrid AI System Delivering Highly Targeted News to Business Professionals

Anupriya Ankolekar[1], David Brunner[1]

[1]*ModuleQ, Inc., 19925 Stevens Creek Blvd. Suite 100, Cupertino, CA 95014, USA*

Abstract

Business professionals need timely, relevant news, but struggle to keep up with the large volume of articles published daily. We present MQ, a hybrid AI news recommender system that uses an explicit model of professionals' commercial relationships to deliver highly targeted news recommendations via a simple chatbot UI. Results from a commercial deployment demonstrate that MQ successfully identifies users' commercial relationships, makes useful recommendations, and drives high and sustained user engagement. We show that domain-specific, knowledge-aware refinements to user modeling and recommendation generation can improve performance.

Keywords

Hybrid AI, Knowledge-aware News Recommenders, Data Fusion, User Work Modeling

1. Introduction

Timely, relevant business news is essential for professionals who develop and manage commercial relationships between organizations. These business-to-business (B2B) professionals include sales executives, account managers, and many senior personnel in financial and professional services firms. News provides crucial context for decision-making and may reveal opportunities for new or expanded commercial engagements. For example, when businesses are acquired or divested, this may influence the prospects of their vendors. However, keeping up with the immense volume, complexity and interdependence of relevant news and information under time pressure is a key trigger for information overload in professionals [1, 2].

Recommender systems [3] are designed to help people find valuable items in vast content. However, main-stream news recommenders (e.g., Apple News, Google News) have failed to reduce news overload for consumers [4] and business professionals. These news recommenders are not targeted enough to deliver valuable recommendations for the highly specific, fast-changing needs of B2B professionals. A recent survey [5] found that only 32% of C-level executives felt that information delivered to decision makers in their organization is relevant and timely. Knowledge about B2B professionals' work and the corporate domain where they operate can help identify useful recommendations for an under-served, but important population.

We present an approach to generate highly targeted news recommendations for B2B professionals by combining data generated by users in the course of their work, domain knowledge, and knowledge-aware recommendation algorithms. Our approach is embodied in the ModuleQ[1] News Recommender (MQ). MQ applies data fusion techniques to learn explicitly structured user profiles dynamically from users' communication data, which provide rich signals about the importance and currency of users' commercial relationships. Extensive domain knowledge is used to identify the business entities involved in these commercial relationships and, together with domain-optimized NLP, to detect the presence of these entities in news articles. A custom, multi-stage recommendation engine generates proactive, precise, explainable recommendations. Results from a commercial deployment in the form of explicit user feedback, user engagement and a user survey show that MQ is effective in delivering relevant news to B2B professionals.

In the following, Section 2 describes requirements for news recommenders for B2B professionals and derives design goals for MQ. Related work is discussed in Section 3, specifically knowledge-aware recommenders and data fusion for modeling a user's key commercial relationships. We present our approach to achieving the design goals (Section 4), in particular deep, domain-tuned modeling and explicit representation of user work driving proactive recommendations delivered via a chatbot in a secure isolated deployment. The MQ system and its components are described in Section 5, with a focus on how the approach manifests in the system. MQ's performance, in terms of user feedback and engagement results, is discussed in Section 6, in particular, the effect of domain modeling on the usefulness of MQ's recommendations. We discuss the implications and limitations of our approach in Section 7, concluding with future work.

2. Design Goals

B2B professionals often have strong financial incentives tied to short-term, customer-related business outcomes. Consequently, they tend to focus intensely on developing customer relationships, while avoiding activities with delayed or ambiguous payoffs such as learning and configuring new tools. Motivated by these characteristics of B2B professionals and the shortcomings of existing news recommenders, we focused on the following design goals:

(G1) Precise targeting: Reduce information overload by recommending only important news that is highly relevant to the user's current needs.

(G2) Minimal user burden: Sustain user engagement by minimizing the need for manual profile configuration and by presenting recommendations within the flow of work, so users need not context-switch into a separate application or web portal.

(G3) Near-instant value: Increase adoption by creating a short, compelling initial experience demonstrating the ability of the system to understand the user's priorities and deliver valuable news recommendations within a few minutes.

(G4) Explainability: Earn user trust trust by explaining the system's recommendations, especially when relevance to a user's relationships is not immediately obvious from an article

---

[1]The authors are founders of ModuleQ Inc. and have led the design and development of the MQ system. MQ has been freely accessible since April 2017 as a chatbot on the Microsoft Teams platform and is also available commercially.

headline. Also, explainability facilitates analysis to further refine the system and supports the discovery and application of domain-specific heuristics.

(G5) Security & compliance: MQ processes users' confidential communication data. Unauthorized access to this information could result in financial or reputational damage, as well as legal liability. Therefore, we designed our system to be highly secure and to facilitate compliance with regulatory regimes such as GDPR [6].

3. Related Work

The vast quantity of news published daily causes significant information overload [4] and people around the world seek more relevant, personalized news [7]. This has motivated long-running research into personalized news recommenders that identify relevant news and filter out the vast majority of irrelevant content [8, 9]. While recommender systems have been successfully applied in many domains (e.g. for finding books, movies, music, consumer purchases), news recommenders face specific challenges stemming from the characteristics of news [10]. Large data volume, unstructured format, short shelf life, high item churn, and changing user interests have been identified as factors that "impede the straightforward application of conventional recommendation algorithms" to news [11].

Knowledge-aware recommenders have been proposed to address challenges of recommending news. New articles often describe events that can be characterized in terms of the named entities involved, e.g., organizations, people, and locations. Knowledge about these entities and their relationships can increase precision, diversity, and explainability. Iana et al. [11] identify six types of knowledge-aware recommenders distinguished by how knowledge is applied and how similarity is measured. Our work extends existing knowledge-aware recommenders by modeling user data to automatically generate up-to-date user profiles, and by matching on a combination of entities and topics to drive recommendations.

For resolution of named entities in news articles, we selected Refinitiv Intelligent Tagging[2] (RIT), an NLP engine trained on business news. RIT identifies business entities with high accuracy and resolves them to known business entities with PermId[3] identifiers. It also identifies topics related to the articles. In addition, Refinitiv maintains knowledge about business entities including revenue, headcount, industry, ownership, executives, directors, and associated financial instruments, which MQ uses as domain knowledge about the corporate world.

3.1. Data fusion for Modeling Commercial Relationships

Commercial relationships are a primary driver of B2B professionals' information needs, so modeling these relationships may improve recommender performance. However, these relationships are complex, dynamic, and ambiguous. No single system captures the state of the relationship as it exists in the mind of the professional, although there are clues in business activity traces strewn across multiple work systems including email, calendar, and customer relationship management (CRM) systems. Data fusion is a framework for modeling complex phenomena

---

[2]https://www.refinitiv.com/en/products/intelligent-tagging-text-analytics
[3]https://permid.org/

from their traces distributed across multiple data streams that may be fragmentary, noisy, and intermittent. The JDL Data Fusion model [12] has been applied and proven in domains as diverse as intelligent transportation [13], disaster mitigation [14], and industry sensor networks[15], and has been revised and extended to many problem domains [16].

There is little prior research applying data fusion to model user work, but other modeling approaches have been used in productivity support tools [17]. IBM's activity-centric computing effort [18] introduced *user activities* as a central concept, inferring them from work data, and developing semantic models and tools for system support [19]. Emails have been recognized as being essential to understanding collaborative work [20], e.g. through person-based methods [21]. Another approach [22] uses clustering over a broader set of user work data (email, calendar, contacts) to create an implicit representation of user work.

User work traces are naturally represented as heterogeneous information networks (HIN) [23]: interconnected information graphs embodying rich structural typing and semantic meaning. Safavi et al. [24] used a graph-based representation of user work data to identify their ongoing activities. MQ similarly represents user work data in a HIN, but applies person-based methods to identify the user's important business relationships.

4. Approach

Our approach to achieve design goals (G1)-(G5) consists of the following: *deep, domain-tuned modeling* combined with an *explicit representation* of user interests (G1-G4), *proactive alerts* via a *chatbot interface* (G2) and *isolated deployment* within the organizational perimeter (G5).

Deep, domain-tuned modeling of user work based on detailed work data (e.g., emails, calendar events, and CRM records) captures important aspects of a B2B professional's work context, including the professional's relationships with other people and organizations, inter-dependencies between these relationships, and changes in relationship structure over time. This enables the creation of rich user interest profiles, providing more precise targeting of recommendations (G1), minimal user burden in creating profiles (G2) and near-instant value (G3).

We apply data fusion techniques [12] to generate an *explicit representation* of the user's portfolio of inter-organizational relationships, first identifying relationships through the people involved, then evaluating the dynamics of each relationship, and eventually the likely impact on the user's commercial objectives. This explicit modeling approach enables easier application of domain-specific heuristics and tuning, and also supports explainability (G4). MQ can explain recommendations in terms of identified relationships, which are more intuitively comprehensible for the user and allow the user to detect inconsistencies and make manual adjustments. Machine learning and NLP methods are used throughout the system, but the focus is always on generating explicit representations that can be reasoned about.

To reduce user burden (G2), MQ uses *proactive alerts* that surface information at appropriate times, without any action by the user. Twice daily updates in the morning and afternoon group together relevant updates. Pre-meeting alerts surface news about specific business relationships ahead of related meetings. Real-time alerts are used for particularly important and relevant news. Since alerts have the potential to be disruptive and add to the user's information overload, precise targeting (G1) is crucial.
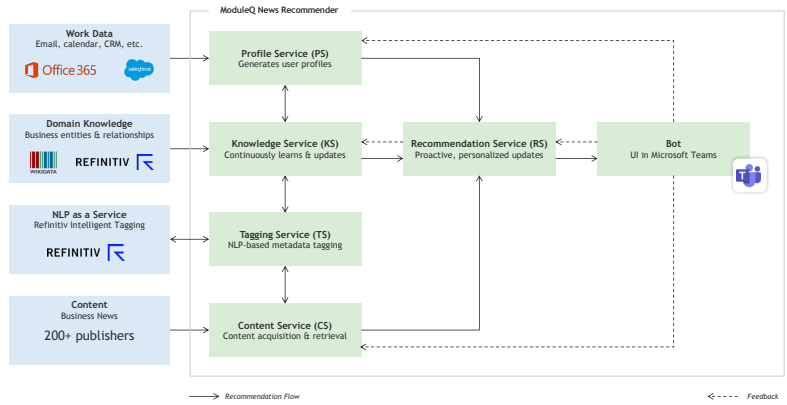
**Figure 1:** Conceptual architecture of MQ, the ModuleQ News Recommender.

To minimize the effort required to access recommendations (G2), MQ provides a *chatbot interface* instead of a standalone application or a web portal. Chatbots interact with users through interactive chat messages, within online workspaces such as Slack and Microsoft Teams that have become widely adopted by professionals. Via the APIs offered by online workspaces, chatbots can deliver alerts within a familiar, cross-platform environment that users access frequently to converse with colleagues throughout their workday. Chatbots support interactive features such as hyperlinks and buttons, simplifying collection of user feedback. Although chatbots are often associated with conversational user interfaces, our system is designed primarily to deliver proactive alerts and does not support conversational interaction.

Finally, for security and compliance (G5), we designed the system for *isolated deployment* within a cloud computing environment controlled by the users' employer. This approach did not materially impact the conceptual operation of the system, but it significantly increased the complexity and cost of implementation, deployment, monitoring, and maintenance. It also limits the volume of training data we are able to gather and thus the usage of data-intensive recommendation methods.

5. ModuleQ News Recommender

Figure 1 depicts the conceptual architecture of the ModuleQ News Recommender (MQ) and its six key subsystems. The *Bot* component includes the chatbot and associated UI in Microsoft Teams. The Bot delivers recommendations, suggests newly identified interests, and collects user behavior data. The *Recommendation Service (RS)* selects content to recommend to each user. The *Knowledge Service (KS)* maintains the system's knowledge about business entities. The remaining three subsystems are primarily concerned with processing inputs to the system: the *Profile Service (PS)* generates user profiles from work data such as emails, calendar meetings, and CRM records; the *Content Service (CS)* ingests and supports retrieval of news articles; and the *Tagging Service (TS)* generates metadata for news articles.

The MQ subsystems were implemented in C# and deployed on Microsoft Azure using Kubernetes. The Bot uses the Microsoft Teams Bot Framework to expose the UI within Microsoft
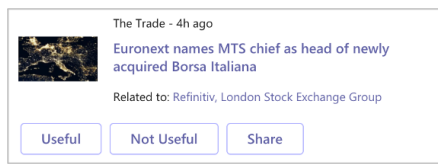
(a) News recommendation.
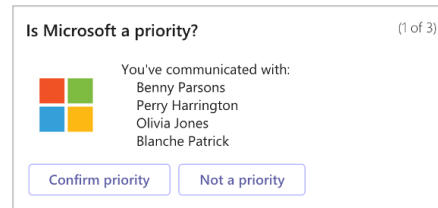
(b) Interest identified by the Profile Service.

Figure 2: Examples of Bot prompts via interactive chat messages in Microsoft Teams.

Teams. For security (G5), processes for developing and operating MQ are certified under the ISO 27001 information security standard. Sensitive user information is encrypted and stays inside the security and governance perimeter of the user's organization.

5.1. Bot and UI

MQ interacts with users via chat messages in Microsoft Teams. Figure 2a shows a news recommendation message displaying the publisher, the article title hyperlinked to the original article, a set of "Related" tags indicating the most significant metadata tags matched with the user's profile, and buttons to provide feedback (*Useful* or *Not Useful*) or to share the article with colleagues. User feedback in the form of clicked hyperlinks and feedback buttons is collected for use in learning.

The Bot also prompts users to provide feedback about the accuracy of commercial relationships identified by PS, as shown in Figure 2b. This feedback is used to refine user profiles, and to analyze the performance of the profile generation. To minimize user burden (G2), the Bot only prompts each user about a small subset of the relationships identified by PS. The full list is made available to the user in a Profile UI, accessible from the chat via a tab in Teams. Revealing the full profile, including unconfirmed interests, supports explainability (G4) and allows users to add or remove interests manually and adjust the recommendations they receive.

5.2. Profile Service

The Profile Service (PS) detects inter-organizational relationships in the user's work data, synthesizing a structured explicit user profile that is used to determine relevant content for the user. Many organizations appear in user data, so PS needs to evaluate which organizations are likely to be of greatest commercial significance to a user at any given time. The profile is updated every hour to ensure an up-to-date profile as the user's relationships evolve (G1). This minimizes user burden (G2) in configuring and maintaining their profiles. PS integrates data from an extensible list of work data sources, currently email, calendar, and CRM records.

PS periodically ingests each user's email messages and calendar meetings. Domain-driven heuristics exclude messages and meetings that are unlikely to be related to significant or current business activity, for instance, emails sent through consumer email providers that are typically not used professionally. PS also ingests recent activities and updates to account, opportunity, and contact records from the user's CRM system. All ingested data is processed to extract

metadata such as email addresses, organizational domains, and timestamps. Email addresses and activity objects are annotated to indicate whether they are external to the user's organization.

Extracted email domains are mapped to the organizational entities that the domains represent with the help of the Knowledge Service (KS). Despite a relatively comprehensive knowledge base of entities, many domains are not resolvable. This is often an indicator that the company is new or too small to be a significant and news-worthy business entity. Nevertheless, we keep track of such domains as they may be important to the user and require inclusion in their profile (G4). This also allows for more accurate estimation of the importance of the relationship, should these domains become associated with more news-worthy entities in the future.

After this preprocessing, PS uses clustering to detect B2B relationships (object assessment in the JDL Data Fusion model) and estimate their salience to the user (situation assessment).

5.2.1. Detecting Relationships with Clustering

Meaningful clusters of business activity are identified in two steps. First, graph-based clustering is used to identify significant activity hot-spots in the user's activity data. A heterogeneous information network (HIN) representing the user's activity data is constructed, where the nodes represent external people (email addresses) or organizations. Edges are created when two people appear together in an activity. A graph-based clustering algorithm then identifies the clusters as connected components in the graph. These clusters typically represent a business relationship with a single external entity; however, due to the complexity of B2B relationships, clusters with multiple organizations are also common.

During the second step, email domains are used to associate clusters with organizations and edges within clusters are weighted by the volume of interaction they represent. The clusters are then refined and pruned by applying domain-specific heuristics. For example, multi-organization clusters may be split if the communication patterns suggest that the cluster really is composed of multiple business relationships. Clusters are pruned if they do not appear to represent a meaningful commercial relationship, e.g. if the communication is only one-way. Accurate knowledge mapping domains to organizations is critical, because many-to-one mappings from domains to organizations are common, and so failure to identify multiple domains as belonging to the same organization can lead to spurious clusters.

5.2.2. Estimating Relationship Salience

For each cluster, we use a set of features to estimate a 'salience' score, i.e. the importance of the corresponding relationship to the user at a given moment in time. For B2B professionals, meetings tend to be particularly significant in predicting the importance of clusters and are hence captured in several hand-engineered features. The features were developed through detailed examination of our own collaboration data and refined through interviews with professionals in customer-facing roles. Additional features used to predict cluster salience include the number of people involved from each organization, the number of interactions between the people, and the proximity of those interactions to the present time. Feature weights were determined offline using training data, collected by prompting users (Figure 2b) to confirm whether organizations associated with the user's highest-salience clusters are important to them in their work.

This process yields a user profile in the form of an ordered set of organizations and people that are important to the user. The user can view the profile created automatically for them and view a summary of evidence (e.g. number of emails, events, people) for why MQ considered a particular organization to be an important business relationship. The user profile is also used to explain MQ's recommendations in terms of the user's relationships.

In addition to tracking organizations and people, the user profile also tracks topics in the form of keyword phrases such as "Cryptocurrencies" or "Tax accounting software". The user profile may be bootstrapped with default topics tuned to the target population for recommendations that are immediately useful (G3). These are added in consultation with the target organization and verified with the user. User topic interests are also learnt automatically based on user feedback on recommendations, and can be entered manually by the user. The universe of topics is learnt from Refinitiv Intelligent Tagging, as described in Section 5.4.

5.3. Content Service

The Content Service (CS) is responsible for ingesting content from external sources and for providing it in clean, annotated form to the rest of the system. In addition to deployment-specific organizational content sources, CS ingests more than 70,000 articles daily from multiple news aggregators. News content contains significant noise in the form of click-bait, automatically generated content, e.g. trading reports, and republished content on multiple sites. CS thus de-duplicates and filters consumer news content, e.g. celebrity news and consumer sales content. Domain-specific rules and heuristics are used to filter the news and to clean the extracted unstructured text. The Tagging Service then annotates the cleaned content with tags and extracts the most reliable and meaningful tags into a content profile. The content along with its profile is saved in a Mongo database, to be used for recommendation and user display. When generating recommendations, the Recommendation Service queries CS for content matching a user profile. CS, in turn, uses Elastic Search for light-weight relevance scoring of articles and returns the top-ranked recent articles as candidates for recommendation. CS also ensures that the candidates represent a balanced distribution of the user's interests.

5.4. Tagging Service

The Tagging Service (TS) aggregates tags from multiple NLP tagging systems, reconciles them and resolves them with the Knowledge Service. For NLP tagging and annotation of content, TS utilizes Refinitiv Intelligent Tagging (RIT) and a custom SpaCy tagging pipeline. RIT provides named entity recognition and annotates content with a rich set of entity and topic tags tailored for business content. Each tag receives a relevance score, representing the relevance of that tag to the article. The organization tags are part of the Refinitiv PermID Linked Data Graph, linking organizations to people, industries and instruments, and are understood by the Knowledge Service. Topics tags are gathered from RIT topic tags and other metadata tags, such as products, technologies, industries and social tags (key noun phrases).

Our custom SpaCy pipeline has been trained for entity recognition specialized to the business domain. News articles often mention additional business entities that are only ancillary to the main story, such as data providers. We developed a custom NLP classifier to annotate mentions

of business entities as data providers, which are subsequently discounted from the entity's relevance score.

At this stage, each article will have multiple, possibly duplicate, tags from different sources. TS clusters these tags, primarily by name, and maps each cluster to an entity within the Knowledge Service. Each tag is also assigned a relevance score that is computed from the source scores. Finally, TS classifies articles and annotates them with article meta-tags, e.g., to identify articles that are likely auto-generated. The list of article tags together with their relevance scores and the article meta-tags forms a content profile.

5.5. Knowledge Service

The Knowledge Service (KS) maintains knowledge about all entities and topics known to the system and provides resolution services to other subsystems. Organizations, people, topics and geographies, as subclasses of the top-level *Thing*. Given an Internet domain or a text string, KS resolves it into an *OrganizationThing*, a *TopicThing* or other appropriate subclasses. Internet domains identified by Profile Service, e.g., are resolved to *OrganizationThing*s in the user's profile. The Tagging Service also annotates content with *Thing*s, enabling the Recommendation Service to match profiles to content. KS can be queried for any *Thing* and returns its model and the metadata around it.

KS currently contains knowledge about 862,798 organizations and 2,145,237 topics. This knowledge is ingested primarily from Refinitiv Content Feeds and the PermID Linked Data graph, but also from Wikidata and directly from organization web pages. For scale, the KS is populated by a mostly-automated knowledge ingestion process, with supplementary manual curation to fix problem data or to handle cases that require human interpretation. It assumes the existence of multiple external sources that may have incomplete data coverage, old data or other data issues. The ingestion process ingests information from the above sources and attempts to reconcile them automatically. If it spots any inconsistencies, these are noted for human review. As domain mapping is a critical part of our system, we have invested significantly into manual curation of organizations with domain mapping. In addition, KS keeps track of the provenance of all knowledge to assist in debugging bad data and diagnosing issues. This also allows for knowledge from any source to be removed if the source is deemed to be of low quality or becomes unavailable for use.

5.6. Recommendation Service

Recommendations are selected from either the news of the past day ($\sim$70K volume) for daily updates or the last two weeks ($\sim$750K) for pre-meeting briefings. The Recommendation Service (RS) queries CS for articles best matching the *OrganizationThing*s and *TopicThing*s in a user profile. RS filters articles that have already been recommended and scores the relevance of remaining articles in multiple steps.

First, the relevance score is computed using modified cosine similarity to consider the salience of a user's interest and the relevance of the interest to the article. High matches on both organizations and topics will result in a high relevance score. Next, references to the user's contacts are identified in the articles, and if present, boost the article's relevance score. For B2B

professionals, any reference to a known client contact is an opportunity to reach out to the client and further the conversation. We reduce noise by only looking for client contacts in articles about the client. Domain knowledge is then used to identify and boost articles with special significance to B2B professionals, such as news about mergers, acquisitions, and leadership changes. The mapping of topics to these significant events is manually curated and maintained in KS. In future work, we plan to boost articles with high predicted engagement based on early user feedback. Finally, article relevance scores are adjusted to prefer recent articles and to prefer articles from authoritative sources.

Having scored and ranked articles, RS selects a small subset for delivery to the user. The selection process picks a diverse set of articles from the top of the ranked list, choosing articles dissimilar to prior recommendations and avoiding too many articles about the same organizations. During scoring, the top *Thing*s (organizations, topics and people) and boosts contributing to the relevance score are captured for display to the user, supporting explainability (G4).

6. Results

This section presents results and findings for an MQ deployment at a global financial technology business with over 10,000 employees. Deployment started with an initial pilot group of about two dozen users in April 2020. Following a successful pilot, the system was rolled out to over 1,000 employees in early 2021. A majority of the users are in customer-facing roles such as sales, account management, and customer success. Besides the quantitative behavioural data and explicit feedback gathered via the Bot, we also collected data on user perceptions via a survey.

6.1. Engagement

Sustained user engagement is an essential performance measure, because busy professionals tend to abandon systems that they do not find useful. Such abandonment shrinks the share of active users relative to the population of registered users. Table 1 shows active and registered MQ users over a six month period. Users are considered active during a time period if they activate the Bot and view one or more messages. 77% of registered users remained active monthly until the end of the six-month period. This metric is conservative, as a significant number of users who left the organization are still counted as registered users. The industry metric DAU/MAU [4] captures how frequently active users use an application. In October 2021, MQ DAU/MAU was 59%, a very high level by industry standards [25].

6.2. Identifying User Interests

In addition to engagement, MQ collects user feedback on how accurately PS identifies organizations that are important to each user. During the initial onboarding process, the Bot prompts users to confirm up to three organizations that were ranked as most salient by PS. The Bot continues to prompt users on an ongoing basis to confirm new organizations that appear highly

---

[4]Daily Active Users (DAU) to Monthly Active Users (MAU) is the average number of users active each day in a given month divided by the number of users active on any day in the month. For our calculation, we exclude weekends when MQ does not deliver recommendations.

Table 1
Registered users and active users per month over a six month period.

	May	Jun	Jul	Aug	Sep	Oct
Registered users	1,043	1,063	1,074	1,131	1,145	1,198
Active users	874	881	850	884	879	919
Active users (%)	84%	83%	79%	78%	77%	77%

(a) User responses to organization priority prompts (n=1160, 279 users, response rate 8.5%).

Response	Onboarding	New Interests
Confirm	206 (74%)	598 (68%)
Reject	72 (26%)	284 (32%)

(b) User-added interests (n=292, 60 users).

Interest Type	Count
Organization	123 (42%)
Topic	169 (58%)

Table 2
User feedback on interests identified by MQ and interests manually added by users.

salient in user work data. For a recent three month period (8th August–16th November 2021), users confirmed 74% of the organizations presented during the initial onboarding stage and 68% of the organizations presented subsequently (Table 2). Users have the option of manually adding organization or topic interests. Only a small number of users (60) chose to add interests, with a slight preference for topics. These results indicate that our approach is effective in determining user interests, in particular organization relationships, from their work data.

6.3. Recommendation Performance

MQ news recommendations are accompanied by *Useful* and *Not Useful* feedback buttons (Figure 2a). Users can click on the headline (*Open*) to read the article on the publisher website. News updates are sent twice daily on weekdays, with each update containing up to four articles.

During the period July 15th–October 14th 2021, MQ delivered about 602,790 news articles sourced from 900 publishers to 1176 registered users. Of these, there were 2395 *Open*s (0.4%) from 472 users. Of the 3231 (0.6%) recommendations with usefulness feedback from 236 users, 59% were marked *Useful* and 41% *Not Useful* (Table 3), equating to a useful ratio[5] of 0.59. This ratio is higher (0.67) for top-ranked articles in daily updates, but decreases to 0.54 for the fourth-ranked articles. The useful ratio is significantly higher for articles matching users' organization interests (0.68) versus those matching only topic interests (0.51), providing evidence that leveraging knowledge about organizations to prioritize recommendations improves performance.

Figure 3 plots the useful ratio observed for article recommendations above a given relevance score percentile. There is a strong positive correlation ($r = .93$, $p < 0.001$) between the relevance percentile and the useful ratio of articles scoring above that percentile. This suggests that the useful ratio could be increased from the 0.6 actually observed up to 0.9 or higher

---

[5] ratio of *Useful* to total feedback. We use this evaluation metric as an indicator of MQ's performance, although it likely incorporates some bias as users choose which articles to give feedback on.

Table 3

User responses to news recommendations (n=3963, 306 users) overall, by rank and by organization/topic matches. The values shown in bold indicate the proportion of Useful in total feedback.

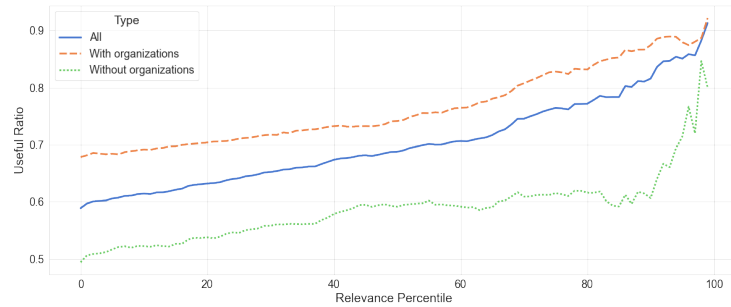| Feedback	All	Rank				Organization matches	
		1	2	3	4	0	1+
Useful	1899 (**0.59**)	575 (**0.67**)	448 (**0.59**)	405 (**0.54**)	428 (**0.54**)	792 (**0.51**)	1125 (**0.68**)
Not Useful	1332	278	309	341	363	774	540
Total	3231	853	757	746	791	1566	1665

Figure 3: Useful ratio for articles with relevance above a given percentile for all articles (solid blue line), articles with Organization tags (dashed orange line), and articles without Organization tags (dotted green line).

by increasing the relevance threshold for recommending news. However, this would reduce recommendation volume and likely recall, potentially lowering overall value to users. We have so far opted to keep recommendation volume, but plan to offer users a choice to receive fewer, higher-relevance recommendations.

We also observe higher useful ratios for recommendations matching users' organization interests. This is consistent across the entire distribution of relevance scores. For recommendations without organization matches, only the top 10% (>90 percentile) reached the level of usefulness achieved by all recommendations with organization matches. These results underline the substantial impact of using a domain-specific knowledge-driven refinement—here, prioritizing users' organization interests—to achieve precise targeting (G1) for a given user population.

6.4. Survey

The user survey was distributed by email in October 2021 to all registered users, of whom 99 users completed the survey (response rate 8.3%). The survey asked users about their perceptions of recommendation volume, timeliness, relevance, and business value. A majority of users responded positively to all of these questions (see Figure 4). Most users agreed that MQ highlights timely (65%) and relevant (61%) insights about the user's interests. This suggests that MQ does not worsen information overload and does deliver relevant news to professional users. Users were ambivalent about whether MQ helped them drive conversations with their clients, a
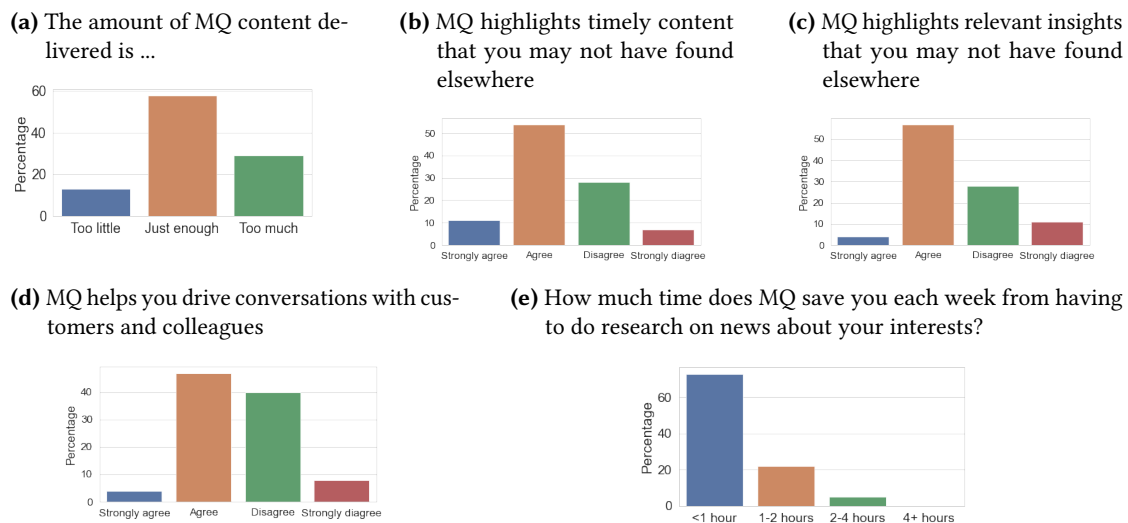
(a) The amount of MQ content delivered is ...

(b) MQ highlights timely content that you may not have found elsewhere

(c) MQ highlights relevant insights that you may not have found elsewhere

(d) MQ helps you drive conversations with customers and colleagues

(e) How much time does MQ save you each week from having to do research on news about your interests?

Figure 4: Summary of user survey responses.

higher bar for usefulness. We also asked users if MQ reduced the amount of time they need to spend searching for information. A significant minority of users (27%) reported substantial time savings, but the majority did not, indicating that proactive recommendations may complement rather than substitute individual research.

7. Discussion

The results above demonstrate the value of a hybrid AI news recommender in supporting the work of B2B professionals with proactive, targeted recommendations. User interviews shed more light on the value of timely news that can be leveraged to further commercial relationships:

> "I was about to have a meeting with [company name], one of the largest energy players in the region ... and received [an MQ card] just before my meeting about latest news on [company name] and their recent expansion plan which I used as a topic of discussion with the client who was very pleased [by] this updated and dynamic interaction. It ... resulted in signing a [dollar amount] deal next day with the client." (user feedback by email, June 22, 2021)

> "[Through MQ] I saw that [company name] had acquired one of the biggest [consumer good] manufacturers in Asia ... I had [a] conversation with the head of trading at [company name] ... from that conversation we've now got a massive integration project going on where we are bringing a load of business from the APAC region for [company name] onto [vendor product] that we didn't see previously ... and that's just from one headline" (user interview, October 7, 2021)

These user stories underline how being aware of corporate developments is important for B2B professionals, and suggest why MQ, being tuned for their work domain, had significant

impact on their work performance. MQ's domain-tuning unfortunately makes it less useful as a news recommender for general-purpose news or for users in other work domains, such as consumer-facing sales or engineering. In general, we expect that high quality recommendations for specific professional populations will likely require specialized recommenders infused with detailed knowledge about the associated professional domain.

Specialized news recommenders may need domain-specific tuning similar to the methods we used to tune MQ for B2B professionals, namely: (1) identify and ingest data that capture particularly important aspects of the professional's work; (2) interpret the data to build a user model that explicitly captures salient aspects of the professional domain, likely with the aid of specialized domain-specific heuristics; (3) annotate content with NLP models trained to identify entities, topics, and events salient to the domain; (4) develop a knowledge base with extensive coverage of entities relevant to the domain, and (5) use knowledge of how different entity categories influence professional work activity to adjust recommendations weights accordingly.

Despite the value of domain knowledge, we faced some pitfalls by assuming it to be relatively stable. Domain knowledge may at times be incomplete, not correspond to user expectations, or become out-dated, which can lead to user confusion or low-quality recommendations in actual use. We encountered user confusion when an organization better known by its pseudonym (sometimes called a "trade name" or "DBA" for "doing business as") was displayed by its legal name in the MQ UI. Another example involved one organization being acquired by another. As the operations of the two formerly separate organizations became deeply intertwined and user data was migrated, the organization affiliations and user email addresses become inconstant and sometimes inconsistent. MQ assumed a single organizational affiliation for each user, resulting in user profiles that were out of alignment with actual conditions. These problems reflect the need to validate knowledge with users and to design for changing knowledge.

Survey results and user interviews also indicated a need for geography-aware recommendations. Survey respondents who felt that MQ delivered too much content were more likely to be located in non-US geographies. MQ allows users to add geographic interests, but this functionality performed poorly because geographies are rarely interests by themselves, functioning instead to moderate the relevance of organizations and topics. Therefore, we believe that future work will need to model geographic interests separately and account for their specific nature when generating recommendations.

The high, sustained user engagement with MQ suggests that a simple chatbot UI may be an effective means of integrating AI into the workflow of busy professionals in many fields. An interesting observation relevant to recommender designers is that common consumer design patterns may not translate well to professional contexts. We experimented with soliciting user feedback with the thumbs-up/thumbs-down, Like/Dislike buttons commonly used in social media, and the Useful/Not Useful buttons described above. We switched to the latter after user engagement increased significantly. Users appear to find news useful without 'liking' it.

The results of our deployment indicate that the users are positive about MQ's recommendations, but that there is scope for them to be more relevant and to help drive customer conversations. In particular, user interviews indicated that explicitly modeling salient aspects of users' professional roles may help tune recommendations. This could be especially effective for specialized professionals focused on specific technologies, regulatory regimes, types of financial instruments or transactions, or categories of risk. Furthermore, we plan to model additional

business-relevant entities and relationships, and experiment with recommendation algorithms that can utilize these effectively in the recommendation process.

8. Conclusion

We presented MQ, a hybrid AI news recommender system for B2B professionals that models users' commercial relationships and delivers proactive, highly targeted news recommendations via a simple chatbot UI. Deployment at a global enterprise shows that MQ is effective at understanding users' commercial relationships and making useful recommendations. Incorporating salient domain entities into modeling of user work and the recommendation process yields significantly more useful recommendations.

Our work is one of the first to apply recommendation systems to improve information overload for B2B professionals, an important but under-served audience. We further identified news recommendations for specific populations of professionals as a promising area for the development of hybrid AI systems, combining knowledge about the user's work and about the domain to generate highly-targeted recommendations with minimal user effort. Characteristics of MQ's knowledge-aware architecture may be adaptable to other domains, namely, applying data fusion techniques to model explicit, time-varying user profiles from multiple noisy data streams and incorporating domain-specific heuristics to prioritize recommendations that are highly salient to the user's work domain.

Acknowledgments

References

[1] P. G. Roetzel, Information overload in the information age, Business research 12 (2019) 479–522.

[2] A. Edmunds, A. Morris, The problem of information overload in business organisations: a review of the literature, International journal of information management 20 (2000) 17–28.

[3] F. Ricci, L. Rokach, B. Shapira, Introduction to recommender systems handbook, in: Recommender systems handbook, Springer, 2011, pp. 1–35.

[4] A. E. Holton, H. I. Chyi, News and the overloaded consumer: Factors influencing information overload among news consumers, Cyberpsychology, behavior, and social networking 15 (2012) 619–624.

[5] CGMA, Joining the dots: Decision making for a new era, https://www.cgma.org/resources/reports/joining-the-dots.html, 2016. Accessed: 2022-0131.

[6] EU, Directive 95/46/ec (general data protection regulation), https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679, 2016. Accessed: 2021-11-24.

[7] R. K. Nielsen, People want personalised recommendations (even as they worry about the consequences), https://www.digitalnewsreport.org/essays/2016/people-want-personalised-recommendations/, 2016. Accessed: 2022-0211.

[8] S. E. Robertson, I. Soboroff, et al., TREC 2002 filtering track report, in: TREC, 3, 2002, p. 5.

[9] J. R. Frank, M. Kleiman-Weiner, D. A. Roberts, E. Voorhees, I. Soboroff, Evaluating stream filtering for entity profile updates in TREC 2012, 2013, 2014, Technical Report, MIT, 2014.

[10] M. Karimi, D. Jannach, M. Jugovac, News recommender systems–survey and roads ahead, Information Processing & Management 54 (2018) 1203–1227.

[11] A. Iana, M. Alam, H. Paulheim, A survey on knowledge-aware news recommender systems, 2021. Under review.

[12] A. N. Steinberg, C. L. Bowman, F. E. White, Revisions to the JDL data fusion model, in: Sensor fusion: Architectures, algorithms, and applications III, volume 3719, International Society for Optics and Photonics, 1999, pp. 430–441.

[13] N.-E. El Faouzi, H. Leung, A. Kurian, Data fusion in intelligent transportation systems: Progress and challenges–a survey, Information Fusion 12 (2011) 4–10.

[14] A. Jotshi, Q. Gong, R. Batta, Dispatching and routing of emergency vehicles in disaster mitigation using data fusion, Socio-Economic Planning Sciences 43 (2009) 1–24.

[15] S. N. Razavi, C. T. Haas, Multisensor data fusion for on-site materials tracking in construction, Automation in construction 19 (2010) 1037–1046.

[16] E. Blasch, A. Steinberg, S. Das, J. Llinas, C. Chong, O. Kessler, E. Waltz, F. White, Revisiting the JDL model for information exploitation, in: Proceedings of Information Fusion, IEEE, 2013, pp. 129–136.

[17] A. Cheyer, J. Park, R. Giuli, IRIS: Integrate, Relate. Infer. Share., Technical Report, SRI, Menlo Park, CA, 2005.

[18] M. Moore, M. Estrada, T. Finley, M. Muller, W. Geyer, Next generation activity-centric computing, in: Proc. CSCW 2006, Citeseer, 2006.

[19] T. P. Moran, A. Cozzi, S. P. Farrell, Unified activity management: supporting people in e-business, Communications of the ACM 48 (2005) 67–70.

[20] V. Bellotti, N. Ducheneaut, M. Howard, I. Smith, Taking email to task: the design and evaluation of a task management centered email tool, in: Proceedings of the SIGCHI conference on Human factors in computing systems, 2003, pp. 345–352.

[21] M. Dredze, T. Lau, N. Kushmerick, Automatically classifying emails into activities, in: Proceedings of Intelligent user interfaces (IUI), 2006, pp. 70–77.

[22] T. M. Mitchell, S. H. Wang, Y. Huang, A. Cheyer, Extracting knowledge about users' activities from raw workstation, in: AAAI, volume 1, 2006, p. 181.

[23] C. Shi, Y. Li, J. Zhang, Y. Sun, P. S. Yu, A survey of heterogeneous information network analysis, IEEE Transactions on Knowledge and Data Engineering 29 (2017) 17–37.

[24] T. Safavi, A. Fourney, R. Sim, M. Juraszek, S. Williams, N. Friend, D. Koutra, P. N. Bennett, Toward activity discovery in the personal web, in: Proceedings of Web Search and Data Mining (WSDM), 2020, pp. 492–500.

[25] A. Chen, DAU/MAU is an important metric to measure engagement, but here's where it fails, https://andrewchen.com/dau-mau-is-an-important-metric-but-heres-where-it-fails/, n.d. Accessed: 2021-11-24.