[illegible] دفاع در مبارزه با تروریسم و [illegible] قانون

مولف: دیویدبی اسکیلکرن

مترجم: مریم ترکاشوند

سرشناسه: اسکیلیکورن، دیوید بی. Skillicorn, David B
عنوان و نام پدیدآور: کشف دانش در مبارزه با تروریسم و اجرای قانون[نویسنده] دیویدبی اسکیلیکورن ؛ مترجم مریم ترکاشوند.
مشخصات نشر: تهران : دانشگاه علوم انتظامی ناجا، ۱۳۹۲.
مشخصات ظاهری: ۴۹۷ص.
شابک: ۶-۴۵-۶۶۴۷-۶۰۰- ۹۷۸
وضعیت فهرست نویسی: فیپا
یادداشت عنوان اصلی: Knowledge discovery for counterterrorism and law enforcement c2009.
موضوع: داده‌کاوی برای اجرای قانون
موضوع: تروریسم -- پیشگیری
شناسه افزوده: ترکاشوند، مریم، ۱۳۵۶ -، مترجم
شناسه افزوده: دانشگاه علوم انتظامی ناجا
رده بندی کنگره: QA۷۶/۹ ۱۳۹۲الف ۲
رده بندی دیویی: ۰۰۶/۳۱۲
شماره کتابشناسی ملی: ۳۳۰۲۰۶۹

انتشارات دانشگاه علوم انتظامی
مولف: دیوید اسکیلیکورن
مترجم: مریم ترکاشوند
ویراستار: فرزین فولادی
صفحه‌آرا: خدیجه کبیری
طراح جلد: محمدرضا مظاهری
چاپ اول: زمستان ۹۲
قیمت: ۲۰۰۰۰ تومان
لیتوگرافی، چاپ و صحافی: مرکز چاپ و نشر عقیدتی سیاسی ناجا (ایمان)

فهرست مطالب

عنوان.....صفحه

مقدمه..... ۱۲

پیش‌گفتار..... ۱۳

فصل اول: کشف دانش..... ۱۷

۱-۱- کشف دانش چیست؟..... ۱۹

۱-۱-۱- حالات اصلی کشف دانش..... ۲۰

۱-۱-۲- فرایند گسترده‌تر..... ۲۲

۱-۲- زمینهٔ تخاصمی چیست؟..... ۲۶

۱-۳- کشف دانش با استفاده از الگوریتم..... ۳۳

۱-۳-۱- چه تفاوتی در مورد کشف دانش در زمینهٔ تخاصمی وجود دارد؟..... ۳۸

۱-۴- آخرین تحولات..... ۴۱

فصل دوم: داده‌ها..... ۴۹

۲-۱- انواع داده‌ها..... ۵۰

۲-۱-۱- داده‌هایی در مورد اشیا یا موضوعات..... ۵۰

۲-۱-۲- داده‌های سطح پایین..... ۵۲

۲-۱-۳- داده‌هایی در مورد ارتباطات..... ۵۳

۲-۱-۴- داده‌های متنی..... ۵۵

۲-۱-۵- داده‌های مکانی..... ۵۶

۲-۲- داده‌هایی که تغییر می‌یابد..... ۵۶

۲-۲-۱- تغییرات تدریجی در شرایط زیربنایی..... ۵۶

۲-۲-۲- تغییر، همان ویژگی حائز اهمیت است..... ۵۷

۲-۲-۳- سیل داده‌ها..... ۵۹

۲-۳- تلفیق انواع مختلف داده‌ها..... ۵۹

۲-۴- چگونه داده‌ها را جمع‌آوری می‌کنند؟..... ۶۴

۲-۴-۱- نقاط پایانی تراکنش..... ۶۴

۲-۴-۲- نقاط پایانی تعامل..... ۶۴

۲-۴-۳- نقاط پایانی مشاهده ۶۵
۲-۴-۴- جمع‌آوری داده به‌وسیلهٔ انسان ۶۶
۲-۴-۵- دلایلی برای جمع‌آوری داده ۶۶
۲-۵- آیا می‌توان به داده‌ها اعتماد کرد؟ ۶۸
۲-۵-۱- مشکل وجود اختلال ۷۱
۲-۶- چه مقدار داده؟ ۷۲
۲-۶-۱- بر هم عمل‌پذیری داده‌ها ۷۴
۲-۶-۲- دانش حوزه ۷۶
فصل سوم: اصول بلند پایه ۷۹
۳-۱- در جست‌وجوی چه باشیم؟ ۸۰
۳-۲- براندازی کلیشهٔ دانش ۹۱
۳-۲-۱- براندازی مرحلهٔ جمع‌آوری داده ۹۱
۳-۲-۲- براندازی مرحلهٔ تحلیل ۹۴
۳-۲-۳- براندازی مرحلهٔ تصمیم‌گیری و عمل ۹۴
۳-۲-۴- مشکل جعل داده‌ها ۹۵
۳-۳- تأثیرات ویژگی‌های مربوط به فناوری ۹۹
۳-۴- معنابخشی و آگاهی موقعیتی ۱۰۴
۳-۴-۱- دیدگاه‌های جهانی ۱۰۷
۳-۵- به حساب‌آوردن زمینهٔ تخاصمی در طول زمان ۱۱۰
۳-۶- آیا کتاب حاضر به متخاصمان نیز کمک می‌کند؟ ۱۱۲
۳-۷- پس حریم خصوصی چه می‌شود؟ ۱۱۴
فصل چهارم: پیش‌بینی مخاطره و کشف بی قاعدگی ۱۲۱
۴-۱- اهداف ۱۲۲
۴-۱-۱- پندارهای نادرست ۱۲۲
۴-۱-۲- مشکل تغییرپذیری انسان ۱۲۲
۴-۱-۳- مشکل دشواری محاسباتی ۱۲۳
۴-۱-۴- مشکل کمیابی ۱۲۳

۴-۱-۵- مشکل پیش‌دستی توجیه‌پذیر ۱۲۵
۴-۱-۶- مشکل تمایل به بازنگری ۱۲۸
۴-۱-۷- اهداف واقعی چه هستند؟ ۱۲۹
۴-۲- خطوط کلی فناوری پیش‌بینی‌کننده ۱۳۴
۴-۲-۱ ساخت پیشگوها ۱۳۴
۴-۲-۲- صفات ۱۳۷
۴-۲-۳- مقادیر گمشده ۱۳۸
۴-۲-۴- دلایلی برای یک پیش‌بینی ۱۳۹
۴-۲-۵- خطاهای پیش‌بینی ۱۳۹
۴-۲-۶- دلایلی برای خطاها ۱۴۱
۴-۲-۷- ربه‌بزنی ۱۴۳
۴-۲-۸- پیش‌بینی با یک اطمینان مرتبط ۱۴۵
۴-۳- فرصت‌های پنهان‌سازی ۱۴۶
۴-۴- فناوری‌ها ۱۴۸
۴-۴-۱- نمودارهای درختی تصمیم ۱۴۸
۴-۴-۲- مجموعه‌هایی از پیشگوها ۱۵۵
۴-۴-۳- جنگل‌های تصادفی ۱۶۴
۴-۴-۴- دستگاه‌های بردار پشتیبان ۱۶۶
۴-۴-۵- شبکه‌های عصبی ۱۷۵
۴-۴-۶- قوانین ۱۷۷
۴-۴-۷- انتخاب صفت ۱۸۰
۴-۴-۸- پیش‌بینی توزیعی ۱۸۴
۴-۴-۹- پیش‌بینی هم‌زیست ۱۸۵
۴-۵- تاکتیک‌ها و فرایند ۱۸۶
۴-۶- توسعه فرایند ۱۹۳
۴-۷- مورد ویژه: در جست‌وجوی نظیر و همانند ۱۹۵
۴-۸- مورد ویژه: در جست‌وجوی نقاط دور افتاده ۱۹۸

۴-۹- مورد ویژه: رتبه‌بندی فراوانی ۲۰۳
۴-۹-۱- رکوردهای رایج ۲۰۴
۴-۹-۲- رکوردهایی که قبلاً مشاهده شده‌اند ۲۰۵
۴-۹-۳- رکوردهایی مشابه با موارد مشاهده شده قبلی ۲۰۸
۴-۹-۴- رکوردهایی با برخی فراوانی‌های دیگر ۲۰۸
۴-۱۰- مورد ویژه: کشف تفاوت ۲۰۹
فصل پنجم: شباهت و خوشه‌بندی ۲۱۷
۵-۱- اهداف ۲۱۹
۵-۲- خطوط کلی فناوری خوشه‌بندی ۲۲۳
۵-۳- فرصت‌های پنهان‌سازی ۲۲۹
۵-۴- فناوری‌ها ۲۳۱
۵-۴-۱- خوشه‌بندی فاصله- محور ۲۳۲
۵-۴-۲- خوشه‌بندی زاویه- محور ۲۳۵
۵-۴-۳- خوشه‌بندی توزیع- محور ۲۳۷
۵-۴-۴- خوشه‌بندی تجزیه- محور ۲۴۰
۵-۴-۵- خوشه‌بندی سلسله‌مراتبی ۲۴۴
۵-۴-۶- خوشه‌بندی دوتایی ۲۴۷
۵-۴-۷- خوشه‌ها و پیش‌بینی ۲۵۱
۵-۴-۸- خوشه‌بندی هم‌زیست ۲۵۲
۵-۵- تاکتیک‌ها و فرایند ۲۵۳
۵-۶- مورد ویژه: در جست‌وجوی نقاط دورافتاده‌ای که مشاهده گشته‌اند ۲۵۵
فصل ششم: نگاهی بر درون گروه‌ها- کشف رابطه ۲۶۱
۶-۱- اهداف ۲۶۳
۶-۲- طرح کلی فناوری کشف رابطه ۲۶۵
۶-۲-۱- نمودارهای واقعی ۲۶۸
۶-۲-۲- انتخاب در نمودارها ۲۷۳
۶-۲-۳- تحلیل نمودارها ۲۷۴

۶-۳- فرصت‌های پنهان‌سازی ۲۷۶
۶-۴- فناوری‌ها ۲۸۰
۶-۴-۱- تحلیل شبکه اجتماعی ۲۸۱
۶-۴-۲- تصویرسازی ۲۸۳
۶-۴-۳- تطبیق‌دهی الگو/ بازیافت اطلاعات ۲۸۹
۶-۴-۴- جست‌وجوی تک- گره‌ای ۲۹۴
۶-۴-۵- کشف ناحیه غیرعادی ۲۹۸
۶-۴-۶- رتبه‌بندی نمودارها ۳۰۰
۶-۴-۷- خوشه‌بندی روی نمودار ۳۰۳
۶-۴-۸- پیش‌بینی ضلع ۳۰۹
۶-۴-۹- کشف ریزساخت خلاف قاعده ۳۱۰
۶-۴-۱۰- نمودارها و پیش‌بینی ۳۱۶
۶-۵- تاکتیک‌ها و فرایند ۳۱۶
فصل هفتم: کشف از میان داده‌های متنی، آشکار ۳۲۷
۷-۱- متن، درحالی‌که وضعیت درونی را آشکار می‌سازد ۳۳۱
۷-۲- اهداف ۳۳۴
۷-۲-۱- یافتن متون مورد نظر ۳۳۴
۷-۲-۲- یافتن محتوا ۳۳۷
۷-۲-۳- یافتن مؤلفان ۳۳۹
۷-۲-۴- یافتن مشخصات مؤلف ۳۴۰
۷-۲-۵- یافتن فرا اطلاعات ۳۴۰
۷-۳- خطوط کلی فناوری تحلیل متنی ۳۴۰
۷-۳-۱- بررسی محتوا ۳۴۳
۷-۳-۲- بررسی مؤلف اصلی ۳۴۷
۷-۳-۳- بررسی فرا اطلاعات ۳۴۸
۷-۴- فرصت‌های پنهان‌سازی ۳۴۹
۷-۵- فناوری‌ها ۳۵۲

۷-۵-۱- جمع‌آوری داده‌های متنی ۳۵۲
۷-۵-۲- استخراج اسناد جالب توجه از میان یک مجموعه بزرگ ۳۵۶
۷-۵-۳- استخراج مدخل‌های اسم‌دار ۳۵۹
۷-۵-۴- استخراج مفاهیم ۳۶۱
۷-۵-۵- استخراج روابط ۳۶۴
۷-۵-۶- استخراج موضوعات ۳۶۵
۷-۵-۷- استخراج روایت و معنابخشی ۳۶۹
۷-۵-۸- خلاصه‌سازی ۳۷۰
۷-۵-۹- استخراج قصد و گرایش ۳۷۲
۷-۵-۱۰- روندهایی در محتوا ۳۷۳
۷-۵-۱۱- استخراج مؤلف از میان گروه‌های نمونه بزرگ ۳۷۴
۷-۵-۱۲- تحلیل مؤلف اصلی از روی گروه‌های نمونه کوچک ۳۷۵
۷-۵-۱۳- کشف ویژگی‌های مؤلف ۳۸۳
۷-۵-۱۴- استخراج ادعاها ۳۸۴
۷-۶- تاکتیک‌ها و فرایند ۳۸۶
۷-۶-۱- ایفای نقش به‌جای مخاصم ۳۸۷
۷-۶-۲- ایفای نقش به‌جای مخاطب ۳۸۹
۷-۶-۳- ادغام داده‌های حاصله از زمینه‌های مختلف ۳۹۱
فصل هشتم: کشف در ارتباطات محرمانه ۳۹۹
۸-۱- تأثیر پیچیده‌سازی ۴۰۲
۸-۲- اهداف ۴۰۳
۸-۳- فرصت‌های پنهان‌سازی ۴۰۳
۸-۳-۱- رمزنویسی ۴۰۴
۸-۳-۲- جایگزین‌سازی واژگانی ۴۰۶
۸-۴- فناوری‌ها ۴۰۸
۸-۴-۱- انتخاب ارتباط جالب توجه ۴۰۸
۸-۴-۲- استخراج محتوا پس از جایگزین‌سازی ۴۲۳

۸-۴-۳- تعیین مؤلف سند پس از جایگزین‌سازی ۴۳۳
۸-۴-۴- فرا اطلاعات پس از جایگزین‌سازی ۴۳۳
۸-۵- تاکتیک‌ها و فرایند ۴۳۴
۸-۵-۱- کاربرد یک فرایند چندوجهی برای انجام انتخاب ۴۳۴
فصل نهم: کشف وضعیت روانی و عاطفی ۴۴۱
۹-۱- تحلیل چارچوب برای کشف مقاصد ۴۴۳
۹-۱-۱- اهداف ۴۴۵
۹-۱-۲- فناوری کشف تحلیل چارچوب ۴۴۶
۹-۱-۳- فرصت‌های پنهان‌سازی ۴۴۸
۹-۱-۴- تاکتیک‌ها و فرایند ۴۴۸
۹-۲- تحلیل احساسات ۴۴۹
۹-۲-۱- اهداف ۴۵۰
۹-۲-۲- فناوری تحلیل احساسات ۴۵۱
۹-۲-۳- فرصت‌های پنهان‌سازی ۴۵۴
۹-۳- استخراج وضعیت روانی ۴۵۵
۹-۳-۱- اهداف ۴۵۶
۹-۳-۲- فناوری استخراج وضعیت ذهنی ۴۵۷
۹-۳-۳- فرصت‌های پنهان‌سازی ۴۶۹
۹-۴- زبان‌شناسی کاربردی نظام‌مند ۴۷۱
۹-۴-۱- مقدمه‌ای بر زبان‌شناسی کاربردی نظام‌مند ۴۷۲
۹-۴-۲- SFL برای کشف قصد ۴۷۷
۹-۴-۳- SFL برای تحلیل احساسات ۴۷۷
۹-۴-۴- SFL برای استخراج وضعیت ذهنی ۴۷۹
کتاب‌نامه ۴۸۵

مقدمه

کسب علم و دانش و تلاش برای بهره‌گیری از علوم جدید و بومی‌سازی آن برای رفع نیازهای علمی، فرهنگی، اجتماعی و اقتصادی به منظور اعتلای اقتدار امت اسلامی از توصیه‌های مهم اسلام است. علوم پلیسی که همگام با پیشرفت بسیاری از علوم محض و تلفیق آنها با یکدیگر، روز به روز نو می‌شود و مرزهایی تازه از دانش می‌سازد، از جمله علومی است که درخت آن در کشورمان در حال رشد و نمو و ثمردادن است.

ترجمهٔ آثار علمی برجسته و معتبر و اشاعهٔ مفاهیم آنها از جمله رهیافت‌های کارآمد و اثربخش است که می‌توان در سایهٔ آن بستری مناسب برای حرکت نخبگان و کارشناسان علوم پلیسی در کشورمان ساخت. بدیهی است تلفیق همگن تجربه‌های کارشناسان امور پلیسی داخلی با علوم و تجربه‌های کارشناسان و خبرگان علوم پلیسی دیگر کشورها باعث پیشبرد دانش پلیسی در کشور خواهد شد. از این رو، بر آن شدیم تا با ترجمهٔ حاضر، گامی هرچند کوچک در این مسیر بلند برداریم. امید است آثاری از این دست بتواند نویسندگان، محققان، کارشناسان و دانش‌پژوهان را برانگیزاند تا تجربه‌ها و توان علمی خود را متبلور کنند وبه رشتهٔ تحریر درآورند.

معاونت پژوهش دانشگاه علوم انتظامی

پیش‌گفتار

انسان‌ها در زمینهٔ دریافت مقادیر بسیار زیادی از داده‌ها و «معنابخشی» به آنها، یعنی کشف محتوای حائز اهمیت داده‌ها و معانی ضمنی‌شان، بسیار ماهرند. به‌عنوان مثال، چشم‌های ما جریان عظیمی از ورودی‌ها را به شکل کوانتوم نوری* دریافت کرده و سیستم بینایی و مغزمان با استفاده از این ورودی‌ها به‌وجود اشیا و ویژگی‌هایشان در دنیای خارجی پی می‌برد. توانایی‌های ما برای معنابخشی به چنین مقادیر عظیمی از داده‌ها به‌منزلهٔ توانایی‌هایی بی‌اختیار و منفعل نیست. ما قادریم تا با تحت کنترل قراردادن حواسمان، انواع خاصی از محرک‌ها را مورد توجه و تفسیر ویژه قرار دهیم. به‌عنوان مثال، اگر در مورد خرید مدل خاصی از خودرو فکر کنیم، به هر کجا که رویم متوجه آن مدل خاص می‌شویم.

مقولهٔ کشف دانش با استفاده از الگوریتم‌های رایانه‌ای به‌نوعی تقلیدی است از این توانایی‌های طبیعی بشر. ما با استفاده از این شیوه بر آنیم تا پس از دریافت مقادیر بسیاری از داده‌ها، با استفاده از محاسبات الگوریتمی به آنها «معنا ببخشیم»، یعنی درست مانند کاری که مغز بشر انجام می‌دهد، محتوای حائز اهمیت درون آنها را استخراج کنیم. نمونه مثال خوبی که در این زمینهٔ می‌توان مطرح کرد، مربوط به داده‌های مرتبط با متون مختلف است. انسان‌ها در زمینهٔ درک محتوای متون ماهرند، به شرط اینکه، متن مورد نظر به شکلی منسجم و منطقی باشد. با این حال، این مهارت‌ها برای شناخت وضعیت روانی مؤلف متن یا اینکه آیا دو متن توسط یک مؤلف نوشته شده‌اند یا خیر، آنچنان که باید و شاید کارآمد نیستند. همان‌طور که خواهید دید، الگوریتم‌ها در زمینهٔ حل این نوع مشکلات مثمرثمر هستند، اما در درک «معنای» متون به‌خوبی مغز بشر عمل نمی‌کنند؛ بنابراین، کشف دانش فراهم‌آورندهٔ امکانات جدیدی است برای استخراج دانش از درون داده‌ها، هم داده‌های عظیمی که انسان‌ها به‌خوبی قادر به تجزیه‌وتحلیل آنها نیستند و هم داده‌هایی که شکل ظاهریشان متناسب با توانایی‌های بشر نیست. اگر کشف الگوریتمی دانش در موازات با توانایی‌های بشری به خدمت گرفته شود، بیشترین کارآمدی را خواهد داشت؛ بدین صورت که در مواردی که

---

*.ذرات بنیادی تشکیل‌دهندهٔ نور را کوانتوم نوری نامند.

انسان‌ها دچار ضعف‌اند کمبود آنها را جبران کرده و در مواردی که عملکرد انسان‌ها بهتر است، به آنها تکیه می‌کند. حوزهٔ کشف الگوریتمیِ دانش که گاهی کشف دانش از میان داده‌ها (KDD)* یا داده‌کاوی نامیده می‌شود، حوزه‌ای به‌سرعت در حال رشد است که با آمار، علم رایانه و روانشناسی هم پوشانی دارد.

موضوع کتاب حاضر، کشف دانش در موقعیت‌هایی است که در آن برخی گروه‌ها مایلند تا داده‌های جمع‌آوری‌شده را دستکاری کرده یا مراحل تحلیل را به نفع خود تغییر دهند تا بدین طریق بتوانند خود یا اعمالشان را مخفی کرده یا نتیجهٔ کشف دانش را به حالاتی متفاوت از نتیجه واقعی‌اش نشان دهند. به‌عنوان مثال، مجرمان هر زمان که بتوانند با انجام کارهای ساده مثل گذاشتن ماسک اسکی برای جلوگیری از ضبط تصویرشان از طریق دوربین‌های مداربسته یا با انجام کارهای پیچیده مثل دست بردن در سیستم حسابداری برای مخفی‌سازی حساب‌سازی‌های مالی‌شان، به‌منظور جلوگیری از برملاشدن فعالیت‌هایشان اقدام می‌کنند. تروریست‌ها در عین حال که برای انجام حمله آماده می‌شوند، اقداماتی را نیز برای مخفی‌سازی خود، اهداف و فعالیت‌هایشان از دید تحلیلگران اطلاعات انجام می‌دهند.

تمامی سیستم‌های کشف دانش دارای کمبود و کاستی‌هایی هستند که ممکن است به‌طور بالقوه مورد سوء استفاده قرار گیرند. حتی مغز بشر نیز دارای چنین ضعف‌هایی است. به‌عنوان مثال، خطای دید نشانگر این است که چگونه گاهی اوقات چیزهایی را می‌بینیم که در واقع وجود ندارند و یک کارت با دو نقطه روی آن می‌تواند برای نشان‌دادن ضعف دیگر ما یعنی داشتن نقاط کور به کار گرفته شود که در این حالت چیزهایی را که وجود دارد، نمی‌توانیم ببینیم. بسیاری از پدیده‌های جهان ما به‌طور روشنی برای بهره‌برداری از شیوه‌ای که مغز تحلیلگر داده‌های ما بدان صورت کار می‌کند، طراحی شده‌اند: یک صفحه تلویزیون LCD تعداد بسیار زیادی از نقاط رنگی را نمایش می‌دهد و ما به طرز شگفت انگیزی قادریم تا از روی این صفحه نمایش دیجیتال اشیا را «ببینیم». علاوه بر آن، این اشیا و تصاویر همگی سه بعدی به نظر می‌رسند، مگر اینکه دیدنمان را متوقف کرده و به چیزی دیگر فکر کنیم. معماران و طراحان معبد

---

*. Knowledge discovery in data (KDD)

پارتنون در آتن از شیوه «دیدن» انسان استفاده کرده‌اند و ستون‌ها را به‌نحوی ساخته‌اند که بلندتر از اندازه واقعیشان به نظر آید.

تمامی سیستم‌های کشف دانش مستعد تحریف و دستکاری‌شدن هستند. در این زمینهٔ تخاصمیِ کشف دانش، طراحان و کاربران سیستم‌های الگوریتمیِ کشف دانش باید در مورد امکان وجود رقیبان احتمالی آگاه بوده و این آگاهی را در نحوهٔ کارکرد سیستم‌ها بگنجانند. این نشانگر آن است که احتمال انجام دستکاری و تحریف به‌طور کلی شیوه کشف دانش را دچار تغییر می‌کند. الگوریتم‌های اصلی کشف دانش نیز در نوع خود نسبتاً متزلزل بوده و مستعد انجام خرابکاری هستند. با این حال، می‌توان پیشرفت‌هایی در دو حوزه‌ایجاد کرد: «سخت‌سازی» الگوریتم‌ها به‌نحوی که احتمال انجام خرابکاری در آنها سخت‌تر شود و گنجاندن آنها درون سیستم‌هایی عظیم تر به‌طوری که حالت دفاع در عمق ایجاد شود. چگونگی انجام این کار، همان موضوع مورد بررسی ما در کتاب حاضر است.

مخاطبان اولیه کتاب حاضر حرفه‌ای‌های حوزه‌های اطلاعات، مبارزه با تروریسم، ضابطین قانون، کشف کلاهبرداری و دیگر امور قضایی هستند که دغدغهٔ اصلی‌اشان اینچنین مقوله‌هایی بوده و با ابزاری کار می‌کنند که به شیوه‌های خاص یا شیوه‌های دیگر دانش را از میان داده‌ها استخراج می‌کنند. سازمان‌ها به‌طور روزافزونی در حال حفاظت از خود در مقابل تخلفات بوده و بنابراین متخصصان حوزهٔ ضد جاسوسی صنعتی که در حال بازرسی و حسابرسی هستند نیز احتمالاً مطالب کتاب حاضر را مفید خواهند یافت.

من در کتاب حاضر پنج هدف اصلی را دنبال می‌نمایم:

- ارائه محتوایی وسیع‌تر برای تجهیزات تک منظوره که به‌منزلهٔ رایج‌ترین حالت از فناوری‌های مورد استفاده امروزی برای کشف دانش هستند. به‌طور مثال امیدوارم که محتویات این کتاب به آن دسته افرادی که از اینچنین تجهیزاتی استفاده می‌کنند کمک کند تا ببینند که چگونه‌این تجهیزات در سیستم‌های بزرگتر به‌کار گرفته می‌شود یا می‌توان آنها را به‌کار گرفت.
- پیشنهاد فرایندی بهتر برای ادغام ابزار کشف دانش. اغلب ابزارهای کشف دانش امروزی تک منظوره بوده یا تنها امکان حالات ویژه‌ای از کشف دانش را فراهم می‌آورند. با این حال،

در یک زمینهٔ تخاصمی، همکوشی بین ابزار با یکدیگر موجب می‌شود که مجموعه کلی به‌وجود آمده بسیار، بسیار قوی‌تر از مجموع هر یک از اعضایش باشد- البته به شرط آنکه‌این اجزا به روشی معقولانه در کنار یکدیگر قرار گیرند. این تمهیدات همواره آشکار نیستند. سعی من بر آن بوده تا ساختارهایی ارائه کنم برای جمع‌آوری ابزار کشف دانش در کنار یکدیگر به روش‌هایی که موجب افزایش هرچه بیشتر میزان قدرت آنها می‌شود.

• تأکید بر انجام احتیاط در خصوص تفسیر نتایج کشف دانش در زمینه‌های خصومت آمیز تا زمانی که ضعف‌های آنها در مقابل دستکاری و تحریف، به‌طور کامل در نظر گرفته شده است. یکی از ویژگی‌های حساس و تعجب آور فناوری کشف دانش فعلی، مربوط است به چگونگی آسیب‌پذیری آن در مقابل اقدامات انجام گرفته از جانب افرادی که خواهان پنهان‌سازی خود و فعالیت‌هایشان هستند. این مسئله به‌اندازه مکفی تشخیص داده نشده و پنجره‌ای تکایی غیر واقعی بر نتایج حاصله از سیستم‌های جعبه سیاه می‌شود.

• فراهم‌کردن پایگاهی برای مباحثی که می‌باید در مورد نقش کشف دانش و جمع‌آوری داده‌های مرتبط با آن به عموم مردم ارائه شود. گفت‌وگوی بین کارکنان سازمان‌های اطلاعاتی و مجری قانون و توده وسیع‌تر مردم از هر دو طرف مأیوس‌کننده به نظر می‌رسد، این تا حدودی به دلیل وجود تفاوت در رویکردها در خصوص مخاطرات و هزینه‌ها بوده؛ اما تا حدودی نیز به دلیل این است که مسئله جمع‌آوری گسترده داده‌ها و تحلیل آن آنچنان که باید و شاید برای مردم روشن نگشته است؛ بلکه همواره این تمایل وجود داشته که جمع‌آوری گسترده داده‌ها به‌عنوان مقوله‌ای بیضرر فرض شود یا به‌عنوان مقوله‌ای کاملاً مخرب، بدون ارائه روشنی از جنبه‌های منفی و مثبت آن- که برخی از آنها وابسته به فناوری‌ها و چگونگی استفاده از آنها است.

• ارائه مروری بر برخی از فناوری‌هایی که در آزمایشگاه موجود بوده و در حال جا به‌جایی از آزمایشگاه‌های تحقیقاتی به حوزه‌ای با کاربرد همه‌جانبه و گسترده بوده، چگونگی کارکردشان و چگونگی کاربرد آنها در کنار شیوه‌های موجود برای کشف دانش. برخی از این اهداف مرتبطند به آن دسته افرادی که ابزار کشف دانش را به‌طور روزانه استفاده می‌کنند. تمامی‌آنها مرتبط اند به آن دسته که در حال طراحی کاربردهای این

فناوری‌ها در آینده، در حوزه‌های فعلی و همچنین در حوزه‌های جدیدی که در آن هنوز کشف دانش رایج نگشته هستند.

این کتاب همچنین برای محققان حوزه‌های هوش مصنوعی؛ آمار، فناوری اطلاعات و حسابداری مفید است. در حوزه‌های فوق رایج است که کشف دانش را برای حوزه اطلاعات و اجرای قانون به‌عنوان کاربرد تخصصی ابزار و الگوریتم‌های اصلی کمتر یا بیشتر در نظر بگیرند. هنوز این مسئله کاملاً روشن نگشته که ماهیت خصومت آمیز این زمینه‌ها تا چه حد موجب تغییر مشکلات می‌شود؛ بنابراین در واقع، تا حد بسیار محدودی از مقدار اصلی کشف دانش موجود را می‌توان به‌طور ایمن برابر دستکاری‌های تعمدی استفاده کرد. نیاز به انجام تحقیقات جدیدی که به این مسئله بپردازد، همراه با انجام ارزیابی‌هایی پیرامون ضعف‌ها و معایب سیستم‌های کشف دانش مورد کاربرد فعلی احساس می‌شود.

کتاب حاضر همچنین مورد توجه اعضای متفکر توده مردم و به‌طور ویژه آن دسته افرادی که تصمیم‌گیری‌هایی در مورد سیاست عمومی در خصوص جمع‌آوری داده‌ها، انجام نظارت، تحلیل داده‌ها و کشف دانش، نه تنها در زمینه‌های خصومت‌آمیز، بلکه همچنین در حوزه‌هایی مثل مدیریت ارتباط با مشتریان، دولت الکترونیک و برنامه‌ریزی اجتماعی و اقتصادی انجام می‌دهند خواهد بود. اگرچه این حوزه‌ها معمولاً به‌عنوان زمینه‌هایی برای انجام اقدامات خصومت‌آمیز در نظر گرفته نمی‌شوند؛ اما در این حوزه‌ها همواره عده‌ای یافت می‌شوند که به دنبال به بازی گرفتن سیستم بوده و بنابراین بذل مقداری توجه به سوی این احتمال معقولانه به نظر می‌رسد؛ همچنین بسیاری افراد نگران افزایش نظارت‌ها و جمع‌آوری داده‌ها از جانب دولت‌ها هستند که به تمایل تروریسم رو به افزایش اسلامی* انجام می‌گیرد. هنگام مواجهه با این وضعیت احتمال هدف واقع شدن فرد به‌سوی دو خطای کاملاً متفاوت بسیار زیاد است: اینکه رویکردی آرمانگرایانه به خود بگیرد (کشف دانش تا هنگامی که به‌اندازه کافی به جمع‌آوری داده‌ها و تحلیل درست آن

---

* م: مرتبط‌کردن فعالیت‌های تروریسم جهانی به مسلمانان و دین اسلام از جمله سیاست‌های دیرینه دشمنان و بالأخص غربیان برای سرپوش گذاشتن بر اعمال جنایتکارانه خود بوده است که در طول تاریخ همواره شاهد آن بوده‌ایم.

می‌پردازیم تروریسم را فرو خواهد نشاند)؛ یا اینکه رویکردی بسته را اتخاذ کند (کشف دانش پایانی است بر داشتن حق مالکیت خصوصی و انجام عادلانه دموکراسی). مقوله‌های سختی در اینجا مطرح است؛ اما، البته واقعیت بسیار پر دامنه‌تر از آنی است که فکرش را می‌کنیم. محتوای کتاب حاضر برای باز کردن جعبه سیاهی که در آن کشف دانش اغلب در این مبحث معرفی شده است کمک کرده و بنابراین کمک خواهد کرد تا به شرح و توضیح موانع و فرصت‌های واقعی نیز بپردازیم.