

مسیرهای جدید، فراروی مرکز تحقیقات کامپیوتری علوم اسلامی

گزارشی از چهاردهمین نشست علوم اسلامی انسانی دیجیتال
با موضوع «رشد علم و اخلاق پژوهش»



به کوشش: هیئت تحریریه فصلنامه ره‌آورد نور

اشاره

چهاردهمین نشست از سلسله نشست‌های علمی علوم اسلامی انسانی دیجیتال، با موضوع «رشد علم و اخلاق پژوهش» در تاریخ ۱۴ اسفندماه ۱۴۰۱ ش در مرکز تحقیقات کامپیوتری علوم اسلامی برگزار گردید. در ابتدای جلسه، آقای دکتر نیکنام به عنوان دبیر نشست، ضمن تبریک اعیاد شعبانیه و خوشامدگویی به میهمانان، ریاست پژوهشگاه علوم و فناوری اطلاعات ایران آقای دکتر محمد حسن‌زاده را معرفی نمود. ایشان، استادتمام علم اطلاعات و بحث دانش‌شناسی و عضو هیئت علمی دانشگاه تربیت مدرس است. در رزومه علمی و سوابق اجرایی دکتر حسن‌زاده آمده است: عضو کمیته علوم انسانی مرکز تحقیقات سیاست علمی کشور، عضو کمیته داوران جشنواره بین‌المللی فارابی، رئیس انجمن ارتقای کتابخانه‌های عمومی ایران، مدیرمسئول و سردبیر و عضو هیئت تحریریه بیش از ده نشریه علمی - پژوهشی، دبیر کمیته برنامه‌ریزی رشته در شورای برنامه‌ریزی آموزش عالی، استاد سرآمد آموزشی دانشگاه تربیت مدرس در سال ۱۴۰۱، پژوهشگر برتر دانشگاه تربیت مدرس در سال ۱۳۹۹ و برگزیده جشنواره بین‌المللی فارابی.

در این نشست همچنین، جناب آقای دکتر فرهاد شیرانی، عضو هیئت علمی پژوهشگاه علوم و فناوری اطلاعات ایران نیز حضور داشت. ایشان دانش‌آموخته دکترای تخصصی مدیریت با گرایش مدیریت دانش از دانشگاه تکنولوژی مالزی است. هم‌اکنون به عنوان استادیار پژوهشکده فناوری اطلاعات در گروه پژوهشی کسب‌وکار الکترونیک در پژوهشگاه علوم و فناوری اطلاعات مشغول به فعالیت است. از جمله سوابق مدیریتی ایشان، عبارت است از: رییس کتابخانه، مدیریت اسناد و اطلاعات، معاون اطلاع‌رسانی پژوهشگاه و رییس مرکز اطلاعات مدارک علمی ایران. ایشان مقالات متعددی در موضوعات مختلف مدیریت اطلاعات و دانش کارآفرینی، دانش آرشیوهای الکترونیکی و کسب‌وکارهای دانش‌بنیان و... دارد.

آنچه در این نوشتار می‌آید، خلاصه‌ای است از مطالب و مباحثی که در این نشست علمی مطرح شده است.

دکتر محمد حسن زاده:

مرکز نور، تسهیل‌گر پژوهش‌های انسانی و اسلامی

آقای دکتر حسن زاده ضمن تشکر بابت برگزاری جلسه، مرکز تحقیقات کامپیوتری علوم اسلامی را به‌مثابه یک تسهیل‌گر در حوزه علوم انسانی اسلامی خواند و فعالیت‌ها و آثار مرکز را بی‌نظیر دانست. ایشان در ادامه گفت:

«این مرکز، به خواست الهی، مسیرهای جدیدی در حوزه‌های نو باز خواهد کرد و قطعاً به عنوان یک پژوهشگر و یک دانشجو برای من و همه پژوهشگران علوم انسانی و اسلامی، این محصولات قابل استفاده بوده و برکات فراوانی داشته است. لازم است جلساتی علمی و بدون تشریفات، راجع به مسائل مشترک گذاشته شود تا گام‌های مثبتی برای پیشبرد علم و فناوری در کشور برداشته شود.»

علم و فناوری، موضوعی راهبردی

ایشان در ادامه، از اهمیت علم و فناوری به عنوان یک موضوع راهبردی گفتند و خاطرنشان کرد:

«امروز هیچ شخصی نمی‌تواند نقش اساسی و تعیین‌کننده علم و فناوری را نفی کند. علم و فناوری، در بین مسئولان ارشد نظام جمهوری اسلامی ایران به عنوان یک حوزه راهبردی مورد توجه قرار دارد و باید گفت در هیچ زمانی، نگاه تکنیکی یا تاکتیکی به علم و فناوری وجود نداشته است. شاهد این مدعا، برنامه‌های توسعه‌ای کشور و سیاست‌های کلی ابلاغی مقام معظم رهبری

است که در همه آنها، علم و فناوری و نوآوری، جزو مهم‌ترین اولویت‌ها قرار دارد. البته این مهم در سایر کشورها نیز همین اندازه در رشد، خوداتکایی و اعتماد به نفس ملی، تعیین‌کننده بوده است.»

دکتر حسن زاده راجع به پیشرفت علم و فناوری در حوزه علوم اسلامی انسانی گفت:

«آنچه امروز مشاهده می‌شود، این است که متخصصان علوم اسلامی در حوزه‌های فنی و فناوری اطلاعات، جزو سردمداران هستند.»

ایشان در ادامه از لزوم حمایت از مرکز تحقیقات کامپیوتری علوم اسلامی گفت و ابراز داشت:

«منظورم از حمایت، صرفاً حمایت بودجه‌ای نیست؛ بلکه باید همه صاحب‌نظران و محققان در حوزه‌های مختلف، به هر نحوی به ارتقای این محصولات کمک کنند، تا این حوزه هویتی، دچار خطا و کندی در حرکت نگردد.»

ایرانداک؛ برنامه‌ها و راهبردها

دکتر حسن زاده در ادامه راجع به دانشگاه علوم و فناوری اطلاعات ایران (ایرانداک) صحبت کرد و گفت:

«این دانشگاه ۵۴ ساله در طول این سالیان، دارای افت‌وخیزهای زیادی بوده است؛ اما امروز در مسیر رشد قرار دارد. اولین پایان‌نامه‌ها و رساله‌هایی که در کشور دفاع شده، از سال ۱۳۲۰ تا الآن، در اینجا ثبت شده است. پایان‌نامه و رساله ایرانیان خارج از کشور نیز در دیتای ما وجود دارد. این، افتخار بسیار بزرگی است که ما الآن می‌توانیم به میلیون‌ها رکورد دسترسی داشته باشیم که نتیجه زحمات همکاران ما در طی این سال‌های متمادی بوده است.»

ایشان در ادامه، از مشقت‌های پیشینان در امر پژوهش گفت؛ به عنوان نمونه، در گذشته برای پیدا کردن یک کتاب، چندین شهر را می‌پیمودند تا به منبعی دست پیدا کنند.

دکتر حسن زاده با اشاره به اهمیت جمع‌آوری پایان‌نامه‌ها به ذکر یک نمونه پرداخت و گفت:

«در یکی از سال‌ها، یکی از بناهای ملی در تهران را با استفاده از دینامیت‌گذاری ستون‌ها می‌خواستند تخریب کنند. به دنبال این بودند که فرو ریختن بنا به شکلی باشد که آسیب به همسایه‌ها وارد نشود. به همین جهت، در جست‌وجوی نقشه این بنا بودند که در نهایت، به ایرانداک می‌رسند. با بررسی نقشه‌ها و طرح ستون‌ها

و فونداسیون، به این نتیجه می‌رسند که این ساختمان، نه تنها نیاز به تخریب ندارد؛ بلکه می‌تواند برای صدها سال دیگر، به حیات خودش ادامه دهد.»

در ادامه، دکتر حسن‌زاده راهبرد ایرانداک را برشمرد و گفت: «هدف ما از ابتدا فراهم نمودن دسترسی به اطلاعات برای پژوهشگران بوده است. در گذشته، طرحی به نام «غدیر» و «امین» اجرا شد که افراد عضو ایرانداک می‌شدند و ایرانداک هزینه اشتراک با کتابخانه‌ها را پرداخت می‌کرد و کتاب‌ها را به پژوهشگران می‌رساند. ایرانداک بابت ارائه اطلاعات، از مخاطبان خود هزینه‌ای دریافت نمی‌کند؛ اما برای نگهداشت دیتا، هزینه جزیی دریافت می‌نماید. دومین راهبرد ما، تأمین منابع علمی برای دانشگاه‌هاست. ایرانداک به عنوان دبیرخانه تأمین منابع علمی کشور، مفتخر است که با وجود تحریم‌ها و محدودیت‌ها، دسترسی به پایگاه‌های اطلاعاتی کل دنیا را برای دانشگاه‌ها فراهم نموده است. سومین راهبرد ما، تنظیم‌گری است که در قالب تدوین استانداردها، برای کتابخانه‌ها، دستورعمل‌ها و شیوه‌نامه‌ها



دکتر حسن‌زاده:

نگاه ما در حوزه اخلاق پژوهش، نگاه هم‌افزایی بوده و نمی‌خواهیم تضاد منافع ایجاد کنیم. بر آنیم تا با همکاری، این آسیب را ریشه کن و به توسعه علم و فناوری و اعتلای نام پژوهشگران و کشور پردازیم. من در این زمینه، نامه‌ای به وزیر محترم ارشاد و مرکز مالکیت معنوی نوشتم که قبل از ممیزی کتاب‌ها، به کمک سامانه همانندجو، از مالکیت فکری آثار باخبر شوید. واقعیت این است که بعضی مواقع، به دلیل نبود مکانیزم نظارتی، این مشکلات تبدیل به یک عارضه می‌شوند



انجام می‌گیرد. شیوه‌نامه ایران، مبتنی بر دیدگاه‌های نرم‌افزاری امروزه با قابلیت انطباق با شیوه‌های نشانی‌دهی طراحی شده و روزآمدسازی می‌شود؛ صدها پروتکل و برنامه‌های مختلفی که در طول زمان در ایرانداک تولید و عرضه شده و قابلیت استفاده در بخش‌های مختلف را نیز دارد؛ به عنوان نمونه، با وجود اینکه حوزه کشاورزی، حوزه تمرکز ایرانداک نیست، اصطلاح‌نامه کشاورزی را ایجاد نمودیم. اصطلاح‌نامه فنی مهندسی و اصطلاح‌نامه آب محیط زیست نیز از این دست هستند. امروز ما مجموعه‌ای از اصطلاح‌نامه‌ها را داریم که برای کتابخانه‌ها و مراکز اطلاعات، به عنوان یک نقطه شروع می‌تواند خوب باشد و در ادامه، می‌خواهیم اینها را به آنتولوژی (هستی‌شناسی) تبدیل کنیم. فرایند تحول یکی از اصطلاح‌نامه‌ها، هفته گذشته شروع شد و ان‌شاءالله بتوانیم در یک زمان مشخص، اصطلاح‌نامه‌ها را به آنتولوژی (هستی‌شناسی) تبدیل نماییم. این کار، برای ما استفاده دوگانه دارد: اول آنکه بتوانیم از آنها در جست‌وجوها و نمایه‌سازی‌ها استفاده کنیم. دوم اینکه دیگران نیز بتوانند از این امکان بهره‌برداری نمایند. به همین منظور، ارتباطات خوبی در این زمینه برقرار کردیم تا مثلاً در حوزه کشاورزی، سازمان تحقیقات کشاورزی و سایر سازمان‌ها از این قابلیت‌ها استفاده نمایند.»

ایشان در ادامه، راجع وضعیت ثبت پایان‌نامه‌ها در سامانه همانندجو گفت:

«پایان‌نامه‌های وزارت علوم و مراکز زیرمجموعه آن، در این سامانه ثبت می‌شود. از مجموعه وزارت بهداشت، حدود ۵ درصد از دانشجویان پایان‌نامه خود را در این سامانه ثبت می‌کنند. ما به دنبال این هستیم که این نهاد به عنوان یک نهاد حاکمیتی ملی، گستره و وسعت کار آن، نزدیک به ۱۰۰ درصد پایان‌نامه‌های کشور باشد. ان‌شاءالله در آینده بتوانیم حوزه‌های علمیه و مراکز دیگری را که در تولید آثار مشارکت دارند، بتوانیم در این مهم همراه کنیم. مأموریت ما به معنای واقعی، این است که این کار در سطح ملی انجام گیرد.»

نکته مهم، این است که هرچقدر در جامعه‌ای کمیت آثار علمی بالا رود، قاعدتاً کیفیت به همان اندازه رشد نمی‌کند. شاید در گام‌های اول، رابطه معکوسی بین کمیت و کیفیت وجود دارد؛ از این جهت که عدم رعایت استانداردهای اخلاق پژوهش در جامعه علمی، به کج‌روی‌ها، کج‌خلقی‌ها و به عبارتی سرقت علمی منجر می‌شود که ممکن است نظام علمی ما را در داخل کشور و نیز در تعامل با نظام‌های بین‌المللی، دچار آسیب کند. من در مصاحبه‌های مختلف عرض کردم که این عارضه در نظام علمی ما، به نسبت تولیدات



علمی، مقدار قابل توجهی نیست. ما بعد از بررسی‌های مختلف متوجه شدیم که علت بخش عمده‌ای از این سوء استفاده یا به عبارتی استفاده‌های نامناسب، ناآگاهی است؛ به عنوان نمونه، گاهی مطلبی را بدون استناد ذکر می‌کنند که به معنای سوء استفاده تلقی می‌شود. انواع مختلف استنادات وجود دارد و ناآگاهی افراد به این موارد، باعث می‌شود که ناخواسته مرتکب بی‌اخلاقی در امر پژوهش باشیم. جهت جلوگیری از این موارد، ما در ایران‌دک کارهایی را شروع کردیم.»

دکتر حسن‌زاده با اشاره به تصویب

قانون مجلس در سال ۱۳۹۶ مبنی بر لزوم ثبت و همانندجویی قرار گرفتن رساله‌ها و پایان‌نامه‌های مراکز علمی کشور گفت:

«پس از تصویب این قانون، همکاران ما در ایران‌دک با درایت، ایثار و تلاش شبانه‌روزی تلاش کردند تا این سامانه را به نتیجه برسانند. در ابتدا ما همانندجویی‌های مختلف به این شکل نداشتیم.

همکاران ما با زحمت فراوان، همراه با عشق و علاقه توانستند این ابزار مورد نیاز برای اجرای قانون را در اختیار دانشگاه‌ها قرار دهند. طبق قانون دانشگاه‌ها موظف‌اند که پایان‌نامه‌ها و رساله‌های آنان را قبل از دفاع، همانندجویی کنند. اگر همانندجویی اتفاق نیفتد، طبق قانون، امکان پرداخت حق‌الزحمه به استاد راهنما، داور یا هر نوع پرداختی مرتبط با این پایان‌نامه، خلاف قانون محسوب می‌شود. علاوه بر بند قانونی، برای ما بسیار مهم بود که از نظر تکنیکال و فنی، ایران‌دک بتواند این کار را به انجام برساند و الحمدلله، این ساماندهی صورت گرفت که آقای دکتر شیرانی نیز در این زمینه توضیحات لازم را خواهند داد. بررسی ما نشان می‌دهد در مدت زمانی که از اجرای قانون می‌گذرد، مشابهت از میانگین بالای ۸۰ درصد به زیر ۳۰ درصد کاهش یافته است. بنابراین، اگر این قانون و ابزار در اختیار ما نبود، امروز ممکن بود که بالای ۸۰ درصد همانندی در تولیدات علمی داشتیم.

در روایتی، از حضرت علی (علیه السلام) نقل شده که ایشان فرموده: «من تساوی یوماه فهو مغبون.» این در پایان‌نامه هم صادق است؛ یعنی یک پایان‌نامه مشترک از دو نفر هم مغبون

است. علاوه بر جلوگیری از سرقت علمی، هدررفت نیروی انسانی نیز با این کار اتفاق نمی‌افتد. امروز مجموعه‌هایی که با ما همکاری نمی‌کنند، معلوم نیست چه میزان همانندی در تولیدات علمی داشته باشند. کمک به توسعه اخلاق پژوهش، افتخاری برای ماست که در ایران‌دک این مهم نهادینه شده است و الحمدلله می‌توانیم ادعا کنیم که پایان‌نامه‌هایی که دفاع شده‌اند، سطحی از نوآوری، مبتنی بر عدم تشابه را دارند. قطعاً خطاها و آسیب‌هایی ممکن است که وجود داشته باشد و ما هم منکر این نیستیم. به هر حال، زمانی که صنعتی شکل بگیرد، مثلاً شغلی به اسم رفع همانندجویی دیده می‌شود، جلوگیری از این موارد، نیازمند تحقیق و توسعه جدی است.»

در ادامه جلسه، دکتر حسن‌زاده اظهار داشت:

«مرکز مدارک علمی ایران، از سال ۱۳۸۸ تا به امروز در قالب پژوهشگاه اداره می‌شود و ایران‌دک، یکی از معاونت‌های پژوهشگاه است. در این بخش، حدود ۱۴۰ نفر کارشناس به صورت ۲۴ ساعته مشغول فعالیت هستند. ما تعطیلی نداریم؛ زیرا ثبت و همانندجویی، پیوسته در سامانه اتفاق می‌افتد.

در حوزه مدیریت اطلاعات علم و فناوری، این افتخار را داریم که اگر اسم کشور ایران در هر رتبه علمی در رسانه‌های دنیا اعلام شود، ما در سامانه خود بلافاصله این مطلب را منعکس می‌کنیم و اعلام می‌نماییم کشور ما در سال‌های قبل هم چه رتبه‌ای داشته است و در بین کشورهای منطقه، حایز چه رتبه‌ای است. بیش از ۹۰ شاخص برای پژوهشگران، مجلات، دانشگاه‌ها و نظام علم

مدرسه می‌شود، از صفر تا صد موضوع را یاد بگیرد. اولین دوره ما، راجع به مالکیت فکری بود که با همکاری مرکز مالکیت فکری قوه قضاییه برگزار شد. در این دوره، شخص از ابتدا تعریف ایده، مالکیت فکری و چگونگی و پایگاه‌های ثبت آن و استانداردهای آن را فرا می‌گرفت. دومین اسکول یا مدرسه، مدرسه اخلاق پژوهش است.»

دکتر حسن‌زاده در ادامه سخنان خود خاطر نشان کرد: «موارد مختلفی از این مدرسه‌ها توسعه پیدا کرده و فرصت خیلی خوبی برای همکاری فراهم شده تا مدرسه‌ای که برای کل کشور می‌تواند مفید باشد، توسط مدرسان شما ایجاد شود و به صورت مشترک، این دوره‌ها به صورت مجازی و حضوری برگزار شود.»

دکتر حسن‌زاده در ادامه، راجع به پژوهشگاه علوم و فناوری اطلاعات گفت:

«ما قلمرو پژوهشی متنوعی در سه پژوهشکده داریم که ۲۹ عضو هیئت علمی با درجه استادیار و دانشیار در گروه‌ها و دپارتمان‌های: اصطلاح‌شناسی، زبان‌شناسی، علم‌سنجی، مدیریت فناوری اطلاعات و غیره مشغول به فعالیت‌اند. پروژه‌های تحقیقاتی زیادی در حوزه‌های تقاضامحور انجام شده است. همچنین، آزمایشگاه‌هایی را هم ایجاد نمودیم. برای راه‌اندازی مرکز کارآفرینی و رشد نیز مجوز اصولی گرفتیم. ایده سرزمین نوآوری در فضای ۲۰ هکتاری، کاری است که با همکاری شهرداری، بنیاد مستضعفان، کمیته امداد و... شروع کردیم. این مجموعه ملی، در سطح و تراز جمهوری اسلامی ایران، در حوزه نوآوری در زمینه علوم و فناوری اطلاعات قرار است ایجاد شود.»

دکتر حسن‌زاده به توضیح در باره روبوداک ایرانداک گفت و اظهار داشت:

«این سیستم، با مخاطب صحبت و جست‌وجو انجام می‌دهد و سپس، سؤال مخاطب را بازبایی می‌کند. هدف ما این است که روبوداک، در مسیر توسعه به عنوان یک رباط سخنگو برای حوزه‌های ذخیره و بازبایی اطلاعات، عمومی‌سازی شود. در سیاست علم و فناوری ما در آزمایشگاه، نرم‌افزار شبیه‌سازی قانون‌گذاری ایجاد شده که شخص، قبل از تصویب قانون، از تبعات آن آگاه شود. این کار، با همکاری مرکز پژوهش‌های مجلس و شورای عالی انقلاب فرهنگی در حال توسعه است. نشریه ما با وجود فارسی بودن، در SCOPOST نمایه شده است. به همین جهت، همایشی تحت عنوان همایش ملی مدیران فناوری اطلاعات در دستور کار داریم که آقای دکتر شیرانی دبیر آن خواهند بود. در این همایش، مرکز تحقیقات کامپیوتری علوم اسلامی نیز

فناوری و نوآوری در رتبه‌های مختلف از قاره اروپا، آمریکا و آسیا به صورت برخط و رایگان در اختیار همه پژوهشگران قرار دارد. افراد می‌توانند این دیتا را در قالب اکسل و... دریافت کنند. دانش ایران در قالب پایگاه‌های SCOPOST یا همان ISI سابق، به صورت مداوم در قالب سامانه ارائه می‌شود. سامانه تأثیر پژوهشی نیز در آینده نزدیک رونمایی می‌شود. در این سامانه، تأثیرات اقتصادی، اجتماعی، سیاسی و غیره نظام علم و فناوری در کشور ما ارائه می‌شود. همچنین، دانشگاه‌ها و پارک‌های علم و فناوری، از منظر ویژگی‌هایی (پدیداری/دیداری) رتبه‌بندی می‌شوند و ما خواهیم دید که پارک‌ها و دانشگاه‌های ما، تا چه اندازه در نظر مردم، صنعتگران و جامعه قرار دارند. این کار، یک رقابت سالم و سازنده‌ای را در بین دانشگاه‌ها و پارک‌ها ایجاد خواهد کرد.

ما بیش از ۲۰۰ دوره آموزشی دارای مجوز از شورای گسترش برگزار کردیم. بعضی از این دوره‌ها، بالای ۱۷۰ مرتبه تکرار شده است. اخیراً رویکرد ما در این زمینه، از دوره منفرد به دوره اسکول یا مدرسه تغییر یافته است. هدف ما این است که شخصی که وارد

دکتر شیرانی:

سامانه همانندجو، یکی از ابزارهایی است که به توسعه اخلاق پژوهش و پیشگیری از تقلب در تهیه آثار علمی کمک می‌کند. این سامانه، بر اساس قانون پیشگیری و مقابله با تقلب در تهیه آثار علمی در پژوهشگاه علوم فناوری اطلاعات ایران تولید و مورد استفاده قرار گرفت. ما دو بستر قانونی برای مقابله با تقلب در تهیه آثار علمی داریم؛ اول، قانون مقابله با تقلب که مجلس شورای اسلامی آن را مصوب نموده و در ادامه، آیین‌نامه اجرایی این قانون که مصوبه هیئت وزیران است



می‌تواند شرکت کند و با توجه به ویژگی‌های محصولات مرکز، این شایستگی وجود دارد که از جوایز ما که فناوری اطلاعات برتر هست، بهره‌مند شوند.»

ایشان در ادامه، از گام‌های توسعه‌ای در زمینه اخلاق پژوهش گفت و بیان داشت:

«نگاه ما در حوزه اخلاق پژوهش، نگاه هم‌افزایی بوده و نمی‌خواهیم تضاد منافع ایجاد کنیم. برآنیم تا با همکاری، این آسیب را ریشه کن و به توسعه علم و فناوری و اعتلای نام پژوهشگران و کشور بپردازیم. من در این زمینه، نامه‌ای به وزیر

محترم ارشاد و مرکز مالکیت معنوی نوشتیم که قبل از میزبانی کتاب‌ها، به کمک سامانه همانندجو، از مالکیت فکری آثار باخبر شوید. واقعیت این است که بعضی مواقع، به دلیل نبود مکانیزم نظارتی، این مشکلات تبدیل به یک عارضه می‌شوند. با مرکز نشر دانشگاهی تفاهم‌نامه امضا کردیم که فایل کل کتاب‌ها را به ما بدهند و ما خدمات همانندجویی به آنها ارائه کنیم. با سازمان سمت و سایر مراکز، در حال رایزنی و عقد تفاهم‌نامه هستیم تا شبکه کاملی را ایجاد کنیم. الحمدلله، با زحماتی که همکاران ما کشیدند، اتصال الگوریتم به مقالات را برقرار کردیم و مقالات به استخر دانش ما اضافه شده است. همچنین، صفحات وب با یک طرح نظام‌مند، امکان همانندجویی در شبکه جهانی اینترنت هم دارد.»

«سامانه همانندجو، یکی از ابزارهایی است که به توسعه اخلاق پژوهش و پیشگیری از تقلب در تهیه آثار علمی کمک می‌کند. این سامانه، بر اساس قانون پیشگیری و مقابله با تقلب در تهیه آثار علمی در پژوهشگاه علوم فناوری اطلاعات ایران تولید و مورد استفاده قرار گرفت. ما دو بستر قانونی برای مقابله با تقلب در تهیه آثار علمی داریم؛ اول، قانون مقابله با تقلب که مجلس شورای اسلامی آن را مصوب نموده و در ادامه، آیین‌نامه اجرایی این قانون که مصوبه هیئت وزیران است. در تبصره ۹ این قانون، به صراحت ذکر شده که حوزه‌های علمیه، دانشگاه‌ها، پژوهشگاه‌ها، مؤسسه‌های آموزش عالی و مجموعه وزارت علوم تحقیقات فناوری وزارت بهداشت، دانشگاه آزاد اسلامی و مجموعه مؤسسه‌های وابسته به دستگاه‌های اجرایی موظف‌اند کلیه مراحل پایان‌نامه یا رساله خود، اعم از پروپزال را قبل از دفاع، همانندجویی کنند و بعد از دفاع، الزاماً در سامانه ثبت نمایند. دانشجویان با ثبت پایان‌نامه خود در ایرانداک، کد ثبتی دریافت می‌کنند. پرداخت مالی استادها را، مشاور و داور و حتی امتیازاتی که در ترفیع و ارتقا مؤثرند، همگی منوط به انجام این کار است.

در ادامه، دکتر حسن‌زاده ابراز امیدواری کرد که: «با توجه پیشرفت‌ها و بسترهای خوبی که مرکز تحقیقات کامپیوتری علوم اسلامی در حوزه علوم و معارف اسلامی ایجاد کرده، بتوانیم با هم‌افزایی کاری بزرگ برای کشور در این حوزه انجام دهیم. ما باید فرصت امروز را غنیمت بشماریم و فکر کنیم در چه مواردی می‌توانستیم همکاری کنیم و این کار را نکرده‌ایم.»

دکتر فرهاد شیرانی:

سامانه همانندجو؛ جایگاه و قابلیت‌ها

در بخش دوم جلسه، دکتر شیرانی راجع به سامانه همانندجو گفت:

با استفاده از زیرساخت‌ها و مستندات خود با مواردی که خارج از فرایندهای معمولی دیتایی را بارگذاری می‌کنند، مواجهه می‌کند. در حال حاضر، ۵۰,۰۰۰ مقاله، ۳۲۰,۰۰۰ پروپوزال و ۶۲۰,۰۰۰ رساله یا پایان‌نامه را زیرمجموعه این بخش داریم که حدود ۱,۴۴۰,۰۰۰ رکورد اطلاعات وجود دارد و ۱۲۴,۰۰۰,۰۰۰ برگ اطلاعاتی، علاوه بر این دو دیتابیس وجود دارد. یکی از دیتابیس‌ها الآن مرتباً به مجموعه اضافه می‌شود؛ یعنی مثلاً اگر دانشجویی فصل اول پایان‌نامه‌اش را نوشت و بارگذاری کرد، بلافاصله به دیتابیس اصلی اضافه شده و بعد از آن، هیچ شخصی نمی‌تواند همان فصل اول را بارگذاری نماید. بنابراین، نفر بعدی که با کسری از ثانیه بارگذاری می‌کند، سامانه اعلام می‌کند با قبلی مشابهت دارد و این، مربوط به فرد قبلی است. بیش از ۸,۰۰۰,۰۰۰ رکورد مقالات لاتین هم آماده شده که امیدواریم به‌زودی در سیستم قرار بگیرد.»

عضویت در سامانه همانندجو

«فرایند عضویت سامانه همانندجو، به این شکل است که ابتدا باید سازمانی که بنیاد کار همانندجویی را انجام دهد، عضو شود. سپس، دانشگاه یا مؤسسه یا ناشر یا برگزارکننده سمینار یا سازمانی که طرح پژوهشی را سفارش می‌دهد، به عنوان کارفرمای طرح پژوهشی در سامانه ثبت‌نام کند. این ثبت‌نام باید از یک طریق شفاف و روشن با نامه کتبی و... اتفاق افتد؛ زیرا نماینده سازمان مذکور در نهایت، مدیریت سامانه همانندجو در آن سازمان را به

عهده می‌گیرد. سپس، فرم‌های مشخصی به صورت آنلاین پُر می‌شود و در آخر، نماینده سازمان، اعضای هیئت علمی را عضو این سیستم می‌کند.

گام دوم این است که عضو هیئت علمی به عنوان یک فرد کامفیرم‌شده و شناخته‌شده در سیستم تلقی شده و فعالیت‌های دانشجویی و سایر فعالیت‌ها در ارتباط مستقیم با عضو هیئت علمی و استاد مورد نظر قرار می‌گیرد. این فرایند تکمیل می‌شود تا اینکه دانشجوی بخواهد عضو شود. عضویت دانشجویی، با دریافت نام کاربری و رمز عبور صورت می‌گیرد. سپس،

ما در سامانه همانندجو، سه بخش بسیار مهم داریم؛ بخش اول، نوشتارهایی است که بارگذاری و ارسال می‌شود و لازم است از کیفیت و استاندارد خاصی برای بارگذاری برخوردار باشند تا فعالیت‌های همانندجو درست انجام بگیرد. بخش دوم، همانندی دقیق است و بخش سوم که واری هوشمند نام‌گذاری شده، این است که تغییراتی را که در متون اتفاق افتاده، کشف می‌کند. واری هوشمند، با استفاده از هوش مصنوعی صورت می‌گیرد. سپس، بعد از تشخیص سیستم، قضاوت سیستم توسط انسان کنترل می‌شود تا اطمینان از درستی آن صورت بگیرد. معماری کلی سامانه، با به‌کارگیری دو سامانه همانندجو و واری هوشمند محقق شده است.

در بخش همانندجویی، فرایندهای پروپوزال‌ها، پایان‌نامه‌ها، رساله‌ها، مقالات، کتب و طرح‌های پژوهشی، در حال حاضر در سامانه موجود است. پایگاه داده‌هایی که این مجموعه را حمایت می‌کند، عبارت‌اند از: پایگاه داده پروپوزال‌ها یا پیشنهادها، پایان‌نامه‌ها و رساله‌ها که به نام «پارسا» شناخته می‌شود. کتب مقاله و سایر اطلاعات موجود در وب نیز دیتایی است که از این مجموعه حمایت می‌نماید. برای انجام این کار، سه زیرساخت در غالب زیرسامانه وجود دارد؛ از جمله مدیریت گواهی‌ها و دسترسی‌ها، بخش احراز هویت و بخشی که آمار توصیفی را می‌سازد و در حقیقت، داشبوردهای بی‌آی که در واقع برای سیاست‌گذار، اطلاعاتی را تولید می‌کند و در نهایت، بخشی که



اعلام می‌کند از کدام مؤسسه یا دانشگاه ثبت‌نام نموده است، با این کار، ارتباط بین دانشجو، مؤسسه و استاد برقرار می‌شود. پس از این، هر فعالیتی که دانشجو انجام دهد، یک نسخه از فعالیت به نماینده دانشگاه و استاد نمایش داده می‌شود و استاد می‌تواند این روند را پیگیری کند.»

فرایند همانندجویی

در ادامه، دکتر شیرانی راجع به فرایند همانندجویی گفت: «کارتابل فرایند همانندجویی با توجه به ویژگی‌های هر کدام از اعضا در اختیار آنها قرار می‌گیرد، هر عضو، کارتابل مخصوص به خود را با تمام اطلاعات مورد نیاز دارد. مشخصات محتوای ارسالی را مشخص می‌نماید و سپس، دیتا را ارسال کرده، در صورت لزوم، هزینه آن را پرداخت می‌کند. پس از این، فرد می‌تواند نتیجه درخواست خود را مشاهده نماید و گواهی آن را دریافت کند. در بخش گزارش تشریحی یا گزارشی که به همراه جزئیات ارائه می‌گردد، مشخص می‌شود که از کدام منبع این مطالب برداشته شده و فهرست منابع را نیز نمایش می‌دهد. ارتباط بین منابع را مشخص نموده، درصد همانندی را به صورت روشن اعلام می‌کند که هر کدام از این منابع، در کدام بخش وجود دارند. در فرایند جست‌وجو، مسئله این است که مثلاً اگر پاراگرافی را در ۱۰ منبع پیدا کنید، سیستم به صورت هوشمند متوجه می‌شود که بیشترین میزان از کدام یک اخذ شده است. بنابراین، سیستم ابتدا منبعی را نمایش می‌دهد که بیشترین مشابهت را داشته است. منابع به صورت تکراری نمایش داده نمی‌شود. هر منبعی یک بار نمایش داده شده و بعد با رنگ دیگری منبع جدید مشخص می‌شود. در نهایت، گواهی صادر می‌شود تا میزان و کیفیت همانندی در اختیار دانشجو قرار گیرد. دانشجو یا استاد، با کلیک روی بخش‌هایی که رنگی هستند، بلافاصله منبع اصلی را می‌بیند. اگر این منبع مجله باشد، متصل به وبگاه مجله می‌شود. اگر منبع، پایان‌نامه باشد، در ایرانداک نمایش داده می‌شود. اگر این منبع کتاب باشد، به کتابخانه ملی متصل شده و مشخصات آن ارائه می‌گردد.»

چالش‌های فرایند همانندجویی

دکتر شیرانی در ادامه اظهار داشت: «روش ما در ایرانداک، برای صدور گواهی، بسیار سخت‌گیرانه است. از نظر ما، مسئولیت کار پایان‌نامه دانشجو، با استاد است و ایرانداک به عنوان ابزار در اختیار استاد گذاشته شده است.»

دکتر شیرانی:

چالش قانونی امروز ما، این است که فعالیت‌هایی که در حوزه‌های علمی انجام می‌شود، در ایرانداک ثبت و همانندجویی نمی‌شود. کمتر از یک درصد وزارت بهداشت، با این سامانه کار می‌کند. دانشگاه آزاد اسلامی، حدود ۳۰ درصد همکاری دارد و ۷۰ درصد از مجموع دانشگاه آزادی اسلامی هم با این سامانه کار نمی‌کند. در مجموعه وزارت علوم، بیش از ۹۷ درصد کار با این سامانه انجام می‌شود

یکی از پیچیدگی‌هایی که در فرایند همانندجویی با آن مواجهیم، چسباندن کلمات به یکدیگر است. وقتی سیستم، این‌همانی را بررسی می‌کند، این واژه را با واژه دیگر مقایسه می‌کند. اگر واژه‌ها در یکی جدا، و در دیگری به هم چسبیده باشد، مفهومی این است که واژه جدیدی تولید شده است. حال اگر تعداد این به‌هم‌چسبیده‌ها در متن زیاد باشد، مفهومی این است که این متن، مثل دیگری نیست؛ درحالی‌که دقیقاً عین هم هستند. در نتیجه، باید سیستم قادر باشد تا به‌هم‌چسبیدگی‌ها را تشخیص دهد. علاوه‌براین، باید قادر باشد تا تکرار محتوا در یک متن را نیز شناسایی کند. اینجا یک تحلیل درون‌متنی صورت می‌گیرد و سیستم می‌گوید از ۲۰۰ صفحه پایان‌نامه، ۱۰ صفحه مطلب بیشتر نیست و مابقی تکرار است. بنابراین، امکان صدور گواهی برای این متن وجود ندارد و به عنوان یک پایان‌نامه تلقی نمی‌شود.

تعبیر و جایگزینی واژه‌های هم‌معنا... به جای یکدیگر نیز دیگر چالش ماست؛ مثلاً به جای «برند»، از واژه جدید «نمانام» استفاده کنند و یا اینکه همه «است»‌ها را جایگزین «بود» می‌کنند و یا

مشابهت زیاد در برخی متون، مورد قبول است. از این رو، ما اینها را به استاد ارجاع می‌دهیم؛ مثلاً حاشیه کتاب عروة الوثقی، شاید کل صفحاتی که نوشته شده، از ۱۰ صفحه تجاوز نکند؛ اما این، اشکالی ندارد. بنابراین، مسئولیت تأیید این متن، به عهده استاد راهنماست. یکی از چالش‌ها در پایان‌نامه‌های حوزوی نیز همین مسئله است. ما در این متون، مواردی که شبیه متون دائمی هستند، نظیر: آیات، روایات، اشعار و... را در یک مرحله کنار می‌گذاریم و بعد مشابهت‌یابی انجام می‌دادیم.»

چشم‌اندازها و راهبردها

دکتر شیرانی در باره چشم‌انداز و راهبردهای مربوط به سامانه همانندجویی در ایراندک گفت:

«در این باره سه مسئله وجود دارد. مسئله اول، توسعه محتوا و پایگاه‌هایی است که آن را توسعه می‌دهد. ما در این بخش، کتاب‌های دانشگاهی را فعلاً در دستور کار داریم و شاید نقطه ضعف ما در زمینه منابع غیردانشگاهی، علوم اسلامی است. تلاش می‌کنیم که این محتوا را توسعه بدهیم. در بخش آکادمیک (علمی)، به صورت خیلی وسیع‌تری به دنبال این توسعه هستیم. بنابراین، بخش مقالات بعد از گذشت ده سال، تقریباً تکمیل شده است. از ابتدای آکادمیک آنچه در دسترس وجود دارد و فایل‌های پی‌دی‌اف از آن تهیه شده، ۱,۳۸۰,۰۰۰ مقاله در بخش دانشگاهی تولید شده که بالغ بر ۱۵۰,۰۰۰ عدد از اینها را ما الآن در اختیار داریم؛ مانند مقالات دانشگاه آزاد، وزارت بهداشت و وزارت علوم. در توسعه بخش کتاب، با ناشران وارد گفت‌وگو شدیم و ناشران بزرگی به مجموعه اضافه شدند. طرح‌های پژوهشی را با وزارت

واژه «قرارداد» را به جای «عقود» می‌گذارند که در این متن در آیه قرآن نیز عبارت «أوفوا قراردادها» دیده می‌شود که کاملاً غلط است.

یکی دیگر از مصادیق این کار، کم و زیاد کردن نقطه‌ها و حروف است؛ مثلاً به جای «نتایج»، «نتایج» می‌نویسند و یا به جای «ر» «ز»، «ص» «ض» و... قرار می‌دهند. در این موارد، سیستم می‌گوید متن مغشوش است و قابلیت صدور گواهی را ندارد. همچنین، گاهی درون متن هیچ کاراکتری آشکارا دیده نمی‌شود؛ ولی با استفاده کاراکترهای پنهان در درون متن که چشم نمی‌بیند، مثلاً بعد از هر کلمه‌ای نقطه گذاشتند. با این کار، سیستم مشابهت کشف نمی‌کند. در این موارد، سیستم، همه کاراکتر «ست‌ها» را کنترل می‌نماید و سپس، اگر کاراکتر «ستی» با کاراکتر «ست» فارسی همخوان نباشد و استفاده شده باشد، آن را استخراج می‌کند. بنابراین، در چنین وضعیتی ما علاوه بر همانندجویی، باید تشخیص دهیم که این متن، غیرقابل استفاده است.»

در ادامه بحث، دکتر شیرانی افزود:

«اگر مشابهت متنی زیر ۳۰ درصد باشد، از نظر کارشناسان ما بررسی می‌شود و اگر بالای ۳۰ درصد بود، باید محتوای آن توسط استاد دیده شود. در صورت تشخیص مشابهت کمتر از ۳۰ درصد، کارشناس مشاهده می‌کند که ماشین اشتباه نکرده باشد؛ مثلاً فرمول‌های شیمی خیلی به هم پیوسته و بزرگ هستند و سیستم تشخیص می‌دهد که این به هم پیوستگی، خیلی زیاد است؛ ولی کارشناسان با توجه به ویژگی فرمول‌های شیمی، این به هم پیوستگی را تأیید می‌کنند. در این سیستم، می‌توان نظر کارشناس را نیز ملاحظه نمود.»

دکتر شیرانی:

یکی از پیچیدگی‌هایی که در فرآیند همانندجویی با آن مواجهیم، چسباندن کلمات به یکدیگر است. وقتی سیستم، این‌همانی را بررسی می‌کند، این واژه را با واژه دیگر مقایسه می‌کند. اگر واژه‌ها در یکی جدا، و در دیگری به هم چسبیده باشد، مفهوش این است که واژه جدیدی تولید شده است. حال اگر تعداد این به هم چسبیده‌ها در متن زیاد باشد، مفهوش این است که این متن، مثل دیگری نیست؛ در حالی که دقیقاً عین هم هستند. در نتیجه، باید سیستم قادر باشد تا به هم چسبیدگی‌ها را تشخیص دهد



علوم گفت‌وگو کردیم که اضافه شوند. بنابراین، توسعه دیتابیس‌های ما افزون بر اضافه کردن بخش لاتین است. نرم‌افزار همانندجو، قابلیت جست‌وجوی لاتین، عربی یا هر زبان دیگری را دارد؛ یعنی محدودیت زبانی در جست‌وجوی همانندی نداریم؛ تنها محدودیت ما، در توسعه دیتابیس‌های پشتیبان است. امکان جست‌وجو در اینترنت را داشتیم؛ اما به دلیل مشکلاتی نظیر ممانعت گوگل، فعلاً این بخش متوقف است. در عین حال، تعداد ۸۰,۰۰۰ صفحه از وب را که اعتبار مناسبی دارند، جمع‌آوری و

افزافه کردیم. تعداد ۸,۰۰۰,۰۰۰ رکورد مقالات لاتین را آماده کردیم که ان‌شاءالله سال آینده، بخش لاتین را با قوت اضافه کنیم. سه ماه پیش، کار دیگری در این زمینه اتفاق افتاد و آن، اضافه شدن همانندجوی کتاب‌ها و طرح‌های پژوهشی بود. توسعه در حوزه نرم‌افزار نیز همان‌طور که جنابعالی واقف هستید، به صورت دائمی صورت می‌گیرد.»

دکتر مهدی نیک‌نام: کاستی‌ها و بایسته‌ها

در ادامه بحث، دبیر این نشست علمی، دکتر مهدی نیک‌نام، معاون فنی مرکز تحقیقات کامپیوتری علوم اسلامی، از ارائه‌دهندگان محترم تقاضا نمود تا با حضور در نشست‌های آتی، بحث‌های نظری بیشتری را در این حوزه مطرح نمایند. آنگاه دبیر نشست، آقای دکتر نیک‌نام اظهار داشت:

«راجع به سرقت علمی در کشور، نیاز به یک طرح بزرگ‌تری داریم؛ از این جهت که مفهوم سرقت علمی، به طور کامل روشن نیست و همان‌طور که در مقالات هم می‌خوانیم، هنوز این موضوع محل بحث است. علاوه بر کپی، پارافریز هم با شرایطی، ممکن است سرقت محسوب شود. در زبان‌ها و رشته‌های مختلف، این مطلب متفاوت است. از این رو، به نظر می‌رسد که چهارچوبی برای این مبحث ترسیم نشده است. به همین جهت، از قانون‌گذار انتظار می‌رود که چارچوب مناسبی آماده‌سازی کند.»

لازم است سامانه‌ای که طراحی شده، بررسی شود که کدام یک از مصادیق سرقت علمی را پوشش نمی‌دهد. تصور اینکه الان با سامانه همانندجو به طور کامل از سرقت علمی جلوگیری می‌کنیم، تصور دقیقی نیست؛ چراکه در برخی موارد، صرفاً دانشجو در نتیجه پایان‌نامه، عدد نادرستی را ادعا می‌کند. یا اینکه مثلاً سامانه ما در مقابل پارافریز چه رفتاری باید داشته باشد؟ راجع به ارجاعات پژوهشگر، چه رفتاری باید صورت گیرد؟ به عنوان مثال، اگر ارجاعات پایان‌نامه مشابهت زیادی با سائر مکتوبات داشت، آیا این هم سرقت علمی محسوب می‌شود؟ به نظر می‌رسد که برای بررسی دقیق‌تر این مباحث، باید یک‌سری نشست‌های علمی داشته باشیم.

اینکه مشابهت در پاراگراف اول متن با آخر متن، مثلاً قسمت چکیده و نتیجه‌گیری که حساسیت بیشتری دارد، متفاوت است، اینها موضوعاتی است که باید تأملاتی راجع به آن صورت بگیرد. به نظر می‌رسد که به‌ازای هر رشته، ما نیاز داریم که از متخصصان و اساتید آن حوزه استفاده کنیم که ویژه رشته تحصیلی خود سرقت علمی را تعریف کنند. ما دو تجربه خوبی که در این زمینه داشتیم، یکی کنفرانس پن بود که در سال ۲۰۱۶ برگزار شد و یکی هم مسابقات شناسایی سرقت علمی بود که در راستای مسابقات ریاتیک در دانشگاه امیرکبیر سال ۱۳۹۴ جهاد دانشگاهی برگزار کرد. بنابراین، به نظر می‌رسد از این دست کنفرانس‌ها و برنامه‌ها، برای رسیدن به این چارچوب لازم باشد.»



دکتر شیرانی:

در بخش همانندجویی، فرایندهای پروپوزال‌ها، پایان‌نامه‌ها، رساله‌ها، مقالات، کتب و طرح‌های پژوهشی، در حال حاضر در سامانه موجود است. پایگاه داده‌هایی که این مجموعه را حمایت می‌کند، عبارت‌اند از: پایگاه داده پروپوزال‌ها یا پیشنهادها، پایان‌نامه‌ها و رساله‌ها که به نام «پارسا» شناخته می‌شود. کتب مقاله و سایر اطلاعات موجود در وب نیز دیتایی است که از این مجموعه حمایت می‌نماید



دکتر حسن زاده:

قانون محوری و ایجاد رقابت

در ادامه و در بخش پایانی این نشست، دکتر حسن زاده خاطرنشان کرد: «دو مسئله وجود دارد؛ اول، بحث قانون سرقت علمی است که قانون‌گذار می‌گوید باید گواهی‌نامه داده شود که سرقت صورت گرفته است یا نه. دوم، بحث در حوزه فناوری است که یک هم‌آوردی‌طلبی، رقابت و پیشرفت باید وجود داشته باشد. این دو مقوله، از یکدیگر جدا هستند. در حوزه اجرای قانون، تا زمانی که قانون تغییر پیدا نکرده، ما موظف به اجرای آن هستیم. به همین جهت، گواهی‌ها را دقیق صادر می‌کنیم؛ یعنی گواهی‌ها به شکلی نباشد که دارای اشکال باشد.

بنده خودم به عنوان عضو هیئت علمی، با اکانتی که دارم، وارد سامانه می‌شوم و گاهی این موارد را چک می‌کنم؛ مثلاً در جلسه دفاع دانشجو، بنده نه به عنوان استاد راهنما و داور، بلکه تنها به عنوان مدیر گروه حضور داشتم. در انتها وقتی همه تأیید کردند،

گفتم: از این دانشجو سؤال دارم. به این دانشجو گفتم: نگاه کن تو چشم من و بگو که این پایان‌نامه را خودت نوشتی یا دیگری نوشته؟ این پایان‌نامه، اصلاً نوشته شده یا نه؟ ایشان گفت: استاد من نمی‌توانم دروغ بگویم. من پایان‌نامه را ننوشتم. ما به خاطر صداقت این دانشجو، او را اخراج نکردیم و به او زمان دادیم تا دوباره کار کند و پایان‌نامه جدید بنویسد. حدود ده روز پیش ایشان دفاع کرد. من در سامانه وضعیت این پایان‌نامه را نگاه کردم. سیستم، بار اول ۹۹ درصد مشابهت نشان داده که در طی چند مرحله، درصد مشابهت به ۴ درصد رسیده بود. خیلی جالب است که سامانه به این ۴ درصد گفته که این متن، غلط یا خطاست (false)؛ ولی کارشناس، گزینه صحیح (true) را انتخاب کرده است. در واقع، سامانه این پایان‌نامه را استاندارد نمی‌دانسته است و کارشناس این را به هر دلیلی تأیید کرده است. ما تلاشمان به عنوان پژوهشگاه، این است که این درصد خطاها، هرچه سخت‌گیرانه‌تر، کمتر شود.

البته همیشه دزد از پلیس جلوتر است؛ یعنی این‌گونه نیست که بگوییم چون پلیس و تکنولوژی‌های نو وجود دارد، هیچ دزدی اتفاق نمی‌افتد؛ اما در حوزه‌ای که مربوط به هم‌آوردی‌ها و رقابت‌هاست، باید خیلی همکاری‌ها را توسعه داد. ممکن است الگوریتمی را اینجا شما طراحی کرده باشید که این الگوریتم بسیار قوی‌تر بوده باشد. چرا از این الگوریتم ما بهره نبریم؟ اصولاً مقوله همانندی چرا نباید به مثابه کسب‌وکار، رشد کند؟

یکی از مشکلات ما این است که در مراحل مختلف، استادان پاسخ سامانه را نگاه نمی‌کنند. اگر استاد نگاه کند که این پایان‌نامه ابتدا ۹۹ درصد مشابهت داشته و سپس ۶۰ درصد و بعد ۴۰ درصد شده، مانع از ادامه این پایان‌نامه می‌شود. اگر صنعت همانندجویی به مثابه یک حوزه کسب‌وکار در کشور رشد کند، به نفع ایراندک، جامعه علمی و قانونی است و در نهایت، به نفع همگان خواهد بود.

ما تلاش کردیم، با وجود اینکه زیاد مورد مراجعه هستیم، سرعت سامانه را به‌روز نگه داریم. یک‌سوم اعضای هیئت علمی ما، در حوزه بازاریابی و روی الگوریتم‌های جدید کار می‌کنند.»

پرسش و پاسخ

در انتهای جلسه، یکی از حضار در مورد تکنیک‌های شباهت‌سنجی سؤالاتی را مطرح کرد و گفت: «در صورتی که تغییر پیشرفته در متن صورت بگیرد، مثلاً



دکتر نیکنام:

راجع به سرقت علمی در کشور، نیاز به یک طرح بزرگ‌تری داریم؛ از این جهت که مفهوم سرقت علمی، به طور کامل روشن نیست و همان‌طور که در مقالات هم می‌خوانیم، هنوز این موضوع محل بحث است. علاوه بر کپی، پارافریز هم با شرایطی، ممکن است سرقت محسوب شود. در زبان‌ها و رشته‌های مختلف، این مطلب متفاوت است. از این رو، به نظر می‌رسد که چهارچوبی برای این مبحث ترسیم نشده است



بین کلمات یک متن جابه‌جایی صورت بگیرد یا فراتر از این پارافریزینگ اتفاق بیفتد، این موارد تا چه اندازه پشتیبانی می‌شود؟ راجع به شباهت معنایی، کاری انجام گرفته است؟ همچنین، راجع به مشابهت‌سنجی‌های بین زبانی در صورتی که تقلب اتفاق بیفتد، مثلاً حجم زیادی از یک مقاله انگلیسی صرفاً بازنویسی و یا برگردان به فارسی شود، آیا شناسایی این موارد، در دستور کار هست؟»

دکتر حسن زاده:

«شناسایی الگوها، حوزه پویایی است؛ به‌خصوص اگر بین زبانی و سیمنتیک باشد، این یک مسئله جدی است و ما قاعدتاً به این سمت حرکت می‌کنیم. ما باید به جایی برسیم که سبک افراد در نگاشتن را الگوسازی کنیم. اگر مثلاً آقای حسن زاده نوشته، بگوییم این سبک، قلم شما نیست. اگر به اینجا برسیم، فوق‌العاده است. راجع به مسائلی که فرمودید، لایه‌های مختلف شناسایی مشابهت در برنامه همانندجو هست؛ ولی در این زمینه‌ها، ادعایی نداریم که

ما همه کار را انجام دادیم. ما در حال یادگیری هستیم و در این مسیر، تلاش پیگیر داریم و تحقیق می‌کنیم و از تجربیات مرکز تحقیقات کامپیوتری علوم اسلامی نیز استفاده خواهیم کرد. من تصمیمی گرفتم که دوستان ما در پژوهشگاه، یک جلسه بازدیدی از اینجا داشته باشند و راجع به بحث‌های فنی با همکاران شما صحبت کنند و این گفت‌وگوها، حتماً برکات خوبی برای ما خواهد داشت.»

دکتر شیرانی:

«اگر شخص در یک پاراگراف، کلمه‌ای را اضافه کند، سیستم به شما خواهد گفت که کل این پاراگراف، نوشته دیگری است و فقط زیر کلمه یا کلماتی که اضافه شده خط می‌کشد. در صورت جابه‌جایی کلمات هم سیستم مشابهت‌سنجی انجام می‌دهد و تشخیص مشابهت می‌دهد؛ اما در بحث بازنویسی، طبق موضع قانون، ما به دنبال همانندجویی دقیق هستیم. از این جهت که چالش‌های مشابهت‌سنجی در متون بازنویسی شده، بسیار زیاد است؛ به عنوان مثال، اگر شخصی، جمله «توانا بود هر که دانا بود» را تبدیل کرد به جمله «دانایی، توانایی است» یا «هر کس داناتر است، تواناتر است»، آیا این کار، کپی کردن محسوب می‌شود؟ اگر بگوییم هر جمله‌ای که مفهوم «توانا بود هر که دانا بود» را القا کند، سرقت علمی محسوب می‌شود، پس چگونه افراد بنویسند؟ به همین جهت، ما مسئولیت این موارد را متوجه استاد راهنما می‌دانیم. بنابراین، در این باره ورود پیدا نمی‌کنیم. اگرچه الگوریتم‌ها طراحی و انجام شده است، اما در عمل، نمی‌توان صرفاً با اینها کار را انجام داد؛ به عنوان مثال، اگر دانشجوی زبان انگلیسی صرفاً ترجمه کرد، با توجه به رشته آن، شاید وظیفه‌اش صرفاً ترجمه باشد.

راجع به سرقت علمی یا مسئله اخلاق پژوهش نیز این نکته لازم است گفته شود این بحث، بسیار گسترده و یک اکوسیستم است که اول، مسئله فرهنگی و سپس، مالکیت فکری و بعد، سایه‌نویسی و مسائل مربوط به آن را شامل می‌شود. دزدیدن ایده نیز نوعی سرقت علمی محسوب می‌شود که با استفاده از نرم‌افزار، قابل تشخیص نیست.» ■