



گزارش رصد فرهنگی (۹۵)

شکاف هنجاری سیاست‌گذاری هوش مصنوعی در ایران؛ تحلیل فرهنگی راهبردی

گزارش پیش‌رو «آشکارسازی» ابعاد غفلت هنجاری در سیاست‌گذاری هوش مصنوعی ایران بررسی کرده و برای این منظور از یک رویکرد تطبیقی انتقادی بهره گرفته تا نشان داده شود که دیگر نظام‌های سیاست‌گذاری چگونه با چالش‌های فرهنگی و ارزشی این فناوری درگیر شده‌اند، این گزارش در شهریور ۱۴۰۴ تولید و منتشر شده است.

با وجود آنکه گفتمان رسمی در ایران بر اهمیت راهبردی هوش مصنوعی تأکید می‌ورزد، رویکرد حاکم، رویکردی چالش‌برانگیز، یک‌جانبه و تقلیل‌گرایان است. استدلال محوری این گزارش آن است که سیاست‌گذاری کشور در این حوزه، تحت سیطره پارادایمی «فنی‌امنیتی» قرار دارد. این پارادایم منحصراً بر اهداف کنترل دولتی، کارآمدی بوروکراتیک و اقتدار ملی متمرکز است. رویکرد مذکور، منجر به پیدایش شکافی «هنجاری فرهنگی» عمیقی شده است که طی آن، ابعاد بنیادین این فناوری در حوزه حقوق بشر، ارزش‌های فرهنگی و عدالت اجتماعی به‌صورت نظام‌مند نادیده گرفته می‌شود.

تحلیل تطبیقی اسناد کلیدی سیاست‌گذاری هوش مصنوعی در ایران با چارچوب‌های مرجع بین‌المللی، الگویی از «غفلت اجرایی» را آشکار می‌سازد. «سند ملی هوش مصنوعی» (۱۴۰۳) بسیاری از اصول هنجاری کانونی مانند «شفافیت»، «توضیح‌پذیری»، «انصاف» و «پاسخگویی» را در سطح ارزشی و اخلاقی به رسمیت شناخته است. (شورای عالی انقلاب فرهنگی، ۱۴۰۳)، با این حال، این اذعان ارزشی به سازوکارهای حقوقی عینی و قابل پیگرد برای شهروندان ترجمه نشده است. این شکاف میان «ارزش اعلام‌شده» و «اقدام ملی»، موجب می‌شود که سازوکار درونی فناوری همچنان به‌مثابه یک «جعبه سیاه» حاکمیتی عمل کند. برای مثال، درحالی‌که اسناد بین‌المللی بر ابزارهای مشخصی مانند «ارزیابی تأثیر الگوریتمی» برای تضمین انصاف تأکید دارند (High-Level Expert Group on AI, ۲۰۱۹)، اسناد ملی با وجود اشاره به «نبود سوگیری» به‌عنوان یک اصل اخلاقی، هیچ راهبرد مشخصی برای شناسایی و مقابله فنی با آن پیش‌بینی نمی‌کنند. این الگو نشان می‌دهد که اگرچه نظام سیاستی به دغدغه‌های هنجاری آگاه است، اما در عمل، این دغدغه‌ها به حقوقی قابل مطالبه برای شهروندان بدل نشده‌اند. بدین معنا که اگر یک سامانه هوشمند در مورد شهروندی تصمیمی منفی بگیرد. (مانند رد یک درخواست)، هیچ مسیر مشخصی برای فهم چرایی آن تصمیم وجود ندارد و اگر عملکرد خطاآمیز یک الگوریتم به فردی آسیب برساند، هیچ فرآیند قانونی روشنی برای اعتراض، اصلاح خطا و جبران خسارت پیش‌بینی نشده است. مهم‌تر از آن، سکوت مطلق اسناد بالادستی در حوزه فضای مجازی و هوش مصنوعی در قبال مفهوم «سوگیری الگوریتمی» است. این موضوع نشان‌دهنده ناتوانی یا نبود تمایل نظام سیاستی به شناسایی و مقابله با ظرفیت این فناوری برای بازتولید و تشدید نابرابری‌های اجتماعی موجود است.

این غفلت هنجاری، پیامدهای ملموسی در عرصه‌های مختلف به همراه داشته است. جنگ ژوئن ۲۰۲۵ میان ایران و اسرائیل به‌مثابه یک نمونه مطالعاتی هشداردهنده عمل کرد. تمرکز انحصاری ایران بر ابعاد فنی‌امنیتی، کشور را در برابر «جنگ روایت‌ها» و کارزارهای اطلاعات نادرست که با هوش مصنوعی تقویت می‌شدند، به شدت آسیب‌پذیر ساخت و به مشروعیت‌زدایی از اقدام‌های ایران در افکار عمومی جهانی دامن زد. در عرصه داخلی نیز، این رویکرد به ظهور یک «اقتصاد گیگ» بی‌ضابطه منجر شده است که در آن، پلتفرم‌های بنام الگوهای جهانی «بهره‌کشی الگوریتمی» از نیروی کار را بدون هیچ‌گونه نظارت هنجاری بازتولید می‌کنند.

تحلیل یافته‌ها نشان می‌دهد، که شکاف هنجاری، محصول منطق ساختاری و درونی پارادایم حاکم است. این گزارش برای تبیین ریشه‌های این منطق، از سه مفهوم کلیدی در نظریه‌های انتقادی فناوری و جامعه‌شناسی سیاسی بهره گرفته است:

«نگاه دولت‌وار» (Scott, ۱۹۹۸): تمایل ذاتی دولت به «خواناسازی» و مدیریت‌پذیر کردن جامعه از طریق فناوری، که طی آن دغدغه‌های پیچیده و «ناخوانا»ی شهروندی از محاسبات سیاستی حذف می‌شوند؛ «راه‌حل‌گرایی فناورانه» (Morozov, ۲۰۱۳): تقلیل مسائل عمیقاً سیاسی و اجتماعی (مانند بحران آب یا فساد) به «ابر چالش‌های» فنی که هوش مصنوعی راه‌حل جادویی آن‌ها پنداشته می‌شود و بدین ترتیب، نیاز به اصلاحات ساختاری دشوار دور زده می‌شود؛

«مصنوع/دست‌­ساخته سیاسی» (Winner, ۱۹۸۰): عینیت یافتن این پارادایم کنترل‌محور در معماری زیرساخت‌های دیجیتال کشور، به‌ویژه «شبکه ملی اطلاعات» که به‌طور ذاتی، نظارت متمرکز را تسهیل و دسترسی آزاد را محدود می‌کند. در این چارچوب، آشفتگی نهادی بر سر «سازمان ملی هوش مصنوعی» و انحلال نهایی آن، نه یک شکست اداری، بلکه تبلور یک نزاع عمیق سیاسی بر سر این است که کدام نهاد، کنترل این ابزار نوین و قدرتمند حکمرانی را در دست خواهد داشت.

هرگونه حرکت به‌سوی توسعه‌ای پایدار و عادلانه، نیازمند گذار از رویکرد فنی-امنیتی به یک رویکرد جامع، انسان‌محور و فرهنگی است. این گزارش سه اقدام مهم را برای تحقق این گذار پیشنهاد می‌کند:

اقدام بنیادین: عملیاتی‌سازی اصول هنجاری مندرج در «سند ملی هوش مصنوعی». این اقدام مستلزم آن است که اصولی چون شفافیت، انصاف و پاسخگویی که اکنون به‌صراحت ذکر شده‌اند (شورای عالی انقلاب فرهنگی، ۱۴۰۳)، از سطح اعلامی فراتر رفته و برای هر یک، نهاد متولی مشخص و شاخص‌های عملکردی عینی تعیین گردد تا این ارزش‌ها به حقوقی قابل مطالبه برای شهروندان بدل شوند. هم‌زمان، تشکیل کارگروهی چندجانبه متشکل از متخصصان فنی، حقوقدانان، دانشجویان و نمایندگان جامعه مدنی برای تدوین پیش‌نویس «منشور ملی اخلاق هوش مصنوعی» الزامی است؛

توسعه نهادی و حقوقی: تصویب قانون «ارزیابی تأثیر الگوریتمی» که نهادهای دولتی را ملزم به ارزیابی پیامدهای اجتماعی و حقوقی سامانه‌های هوشمند پرخطر پیش از پیاده‌سازی نماید. به‌موازات آن، تأسیس «دیده‌بان مستقل هوش مصنوعی» برای انجام بازرسی‌های مستقل و ادواری بر الگوریتم‌های بخش دولتی و انتشار عمومی یافته‌ها جهت ایجاد شفافیت و اعتماد عمومی؛

توانمندسازی اجتماعی: سرمایه‌گذاری راهبردی در حوزه «سواد انتقادی هوش مصنوعی» در نظام آموزشی و رسانه‌های عمومی. هدف، توانمندسازی شهروندان برای فهم و به چالش کشیدن ادعاهای فنی است تا از جایگاه «سوژه‌های منفعل» فناوری به «مشارکت‌کنندگان فعال» در راهبری آن بدل شوند.

تداوم مسیر کنونی نه تنها به تعمیق نابرابری‌های اجتماعی، فرسایش اعتماد عمومی و شکل‌گیری یک «حاکمیت الگوریتمی» غیر پاسخگو در حوزه هوش مصنوعی منجر خواهد شد، بلکه آسیب‌پذیری‌های راهبردی کشور را در عرصه‌های نوین قدرت تشدید می‌کند. پیشنهاد‌های ارائه‌شده در این گزارش، مسیری عمل‌گرایانه و ایجابی را برای توسعه هوش مصنوعی در ایران ترسیم می‌کند. مسیری که در آن پیشرفت فناورانه با بلوغ اخلاقی، عدالت اجتماعی

و تعالی فرهنگی اجتماعی بلندمدت ملت ایران همراستا باشد