

۲۴۰۷۸۱

مجموعه مقالات اخلاق در اطلاعات و هوش مصنوعی

اخلاق، حکمرانی و سیاست‌ها (چارچوب‌ها)

و

از چستی به چگونگی

لوجیانو فلوریدی

شهریانو بهلوان مازندرانی

www.ketab.ir

انتشارات آزاد مهر

عنوان و نام پدیدآور

: مجموعه مقالات اخلاق در اطلاعات و هوش مصنوعی /

[ویراستار] لوچیانو فلوریدي؛

Floridi, Luciano

شهربانو پهلوان مازندرانی ۱۳۳۹

: تهران: آزادمهر، ۰۵.۱۴۰۲-۳۰-۷۸۸۳-۶۷۸-۶۲۲-۹۷۸

مشخصات نشر

: کتاب حاضر ترجمه‌ی بخشی از کتاب "Ethics in

یادداشت

information and artificial intelligence

[2021] است.

: هوش مصنوعی -- مقاله‌ها و خطابه‌ها

موضوع

\*Addresses, Artificial intelligence, essays, lectures

هوش مصنوعی -- جنبه‌های اخلاقی -- مقاله‌ها و

خطابه‌ها

Moral and Artificial intelligence -- Addressses, essays, lectures  
ethical aspects -- Addressses, essays, lectures

- تکنولوژی اطلاعات و ارتباطات -- جنبه‌های اخلاقی -

- مقاله‌ها و خطابه‌ها

Information and communications \*

technologies -- Moral and ethical aspects

-- Addressses, essays, lectures

: ۵/۱۵۸ دیویی ۴۸۳۳/۳۰۳

رده بندی کنکوره

: ۹۵۳۱۹۵۵

شماره کتابشناسی ملی

۰۲۱۸۸۸۴۸۶۱۶

انتشارات آزادمهر

نام کتاب: اخلاق در اطلاعات و هوش مصنوعی

نام نویسنده: لوچیانو فلوریدي

نام مترجم: شهربانو پهلوان مازندرانی

شابک: ۰۵-۳۰-۷۸۸۳-۶۷۸-۶۲۲-۹۷۸

نوبت چاپ: اول ۱۴۰۳

شمارگان: ۱۰۰۰

قیمت: ۱۰۴۰۰۰ تومان

آدرس: تهران، خیابان کریم خان، خ میرزای شیرازی کوچه چهار پلاک ۱۶

## فهرست

|       |                                                                        |        |
|-------|------------------------------------------------------------------------|--------|
| ..... | سخن مترجم                                                              | ۷      |
| ..... | مقاله اول: اخلاق، حکمرانی و سیاست‌ها                                   | ۹      |
| ..... | در هوش مصنوعی                                                          | ۹      |
| ..... | فصل یک:                                                                | ۱۱     |
| ..... | مقدمه                                                                  | ۱۱     |
| ..... | اهمیت یک رویکرد اخلاقی                                                 | ۱۱     |
| ..... | اولین رویکرد در توسعه هوش مصنوعی                                       | ۱۱     |
| ..... | فصل دو:                                                                | ۱۸     |
| ..... | چارچوب یکپارچه از پنج اصل برای هوش مصنوعی در جامعه                     | ۱۸     |
| ..... | ۱.۲ مقدمه                                                              | ۱۹     |
| ..... | ۲.۲ هوش مصنوعی                                                         | ۲۱     |
| ..... | ۳.۲ چارچوب یکپارچه از پنج اصل برای هوش مصنوعی اخلاقی                   | ۲۳     |
| ..... | ۱.۳.۲ سودمندی: اوقای رفاه، حفظ کرامت و حفظ سیاره                       | ۲۷     |
| ..... | ۲.۳.۲ عدم سوء استفاده: حریم خصوصی، امنیت و "قابلیت پیش‌بینی"           | ۲۹     |
| ..... | ۳.۳.۲ خودمختاری یا استقلال: قدرت تصمیم‌گیری (تصمیم‌گیری)               | ۳۰     |
| ..... | ۴.۳.۲ عدالت: ترویج رفاه، حفظ همبستگی، پرهیز از بی‌عدالتی               | ۳۲     |
| ..... | ۵.۳.۲ توضیح‌پذیری: توانمندسازی سایر اصول از طریق قابلیت فهم و پاسخگویی | ۳۳     |
| ..... | ۶.۳.۲ یک نمای مختصر (synoptic) خلاصه                                   | ۳۴     |
| ..... | ۴.۲ اخلاق هوش مصنوعی: از کجا و برای چه کسی؟                            | ۳۵     |
| ..... | ۵.۲ نتیجه‌گیری: از اصول تا عمل                                         | ۳۷     |
| ..... | مقاله دوم: از چیستی به چگونگی                                          | ۳۹     |
| ..... | ابزارهای اخلاق هوش مصنوعی، روش‌ها و تحقیق برای تفسیر اصول              | در عمل |
| ..... |                                                                        | ۴۱     |

|    |                                       |
|----|---------------------------------------|
| ۴۲ | چکیده .....                           |
| ۴۴ | مقدمه .....                           |
| ۵۰ | حرکت از اصول به عمل .....             |
| ۵۷ | روش شناسی .....                       |
| ۶۵ | چارچوب نتایج .....                    |
| ۶۸ | بحث در مورد نتایج اولیه .....         |
| ۶۹ | توضیح پذیری به عنوان اصل فراگیر ..... |
| ۷۶ | تمرکز فردی .....                      |
| ۸۰ | راه رو به جلو .....                   |
| ۸۶ | محدودیت‌ها .....                      |
| ۸۹ | نتیجه .....                           |
| ۹۳ | نمایه .....                           |

## سخن مترجم

این کتاب شامل دو مقاله در زمینه اخلاق اطلاعات و هوش مصنوعی از پروفسور لوچیانو فلوریدی است. مقاله اول اخلاق، حکمرانی و سیاست‌ها به چارچوب‌ها و سیاست‌هایی که در اخلاق هوش مصنوعی بحث شده است می‌پردازد و مقاله دوم از چستی به چگونگی راه حل عملی برای ایجاد اخلاق در داخل (در الگوریتم) هوش مصنوعی و نتایج آن می‌پردازد.

لوچیانو فلوریدی فیلسوف و دانشمند اطلاعات ایتالیایی است. او استاد فلسفه و اخلاق اطلاعات در دانشگاه آکسفورد و مدیر آزمایشگاه اخلاق دیجیتال در آمریکا است. فلوریدی یکی از کارشناسان برجسته اخلاق اطلاعات است و به طور گسترده در مورد این موضوع نوشته است. کار او بر پیامدهای اخلاقی فناوری‌های جدید، مانند هوش مصنوعی، داده‌های بزرگ و رسانه‌های اجتماعی متمرکز است.

در آخرین مقاله خود، "اخلاق اطلاعات: یک چشم انداز جدید"، فلوریدی استدلال می‌کند که اطلاعات نه تنها یک منبع ارزشمند، بلکه یک ارزش اخلاقی نیز است. او چارچوب اخلاقی جدیدی برای اطلاعات پیشنهاد می‌کند که آن را "اخلاق اطلاعاتی" می‌نامد.

چارچوب اخلاق اطلاعاتی فلوریدی بر چهار اصل استوار است:

- اصل سودمندی: اطلاعات باید برای سود بردن افراد و جامعه به طور کلی استفاده شود.

- اصل عدم آسیب: اطلاعات نباید برای آسیب رساندن به افراد یا جامعه استفاده شود.

- اصل خودمختاری یا استقلال: افراد باید حق کنترل اطلاعات خود را داشته باشند.

- اصل عدالت: اطلاعات باید به طور عادلانه و منصفانه توزیع شود.

کار فلوریدی در زمینه اخلاق اطلاعات مهم است زیرا رویکرد جدیدی برای تفکر در مورد پیامدهای اخلاقی فناوری‌های جدید ارائه می‌کند. همچنین چارچوبی برای تدوین دستورالعمل‌های اخلاقی برای استفاده از اطلاعات ارائه می‌دهد.

در اینجا خلاصه‌ای از برخی از نکات کلیدی آخرین مقاله فلوریدی آمده است:

- اطلاعات یک ارزش اخلاقی است.

- اطلاعات باید برای سود بردن افراد و جامعه به طور کلی استفاده شود.

- اطلاعات نباید برای آسیب رساندن به افراد یا جامعه استفاده شود.

- افراد باید حق کنترل اطلاعات خود را داشته باشند.

- اطلاعات باید به طور عادلانه و منصفانه توزیع شود.

- چارچوب اخلاق اطلاعاتی فلوریدی رویکرد جدیدی برای تفکر در مورد پیامدهای اخلاقی فناوری‌های جدید است.

- کار فلوریدی چارچوبی برای تدوین دستورالعمل‌های اخلاقی برای استفاده از اطلاعات ارائه می‌دهد.