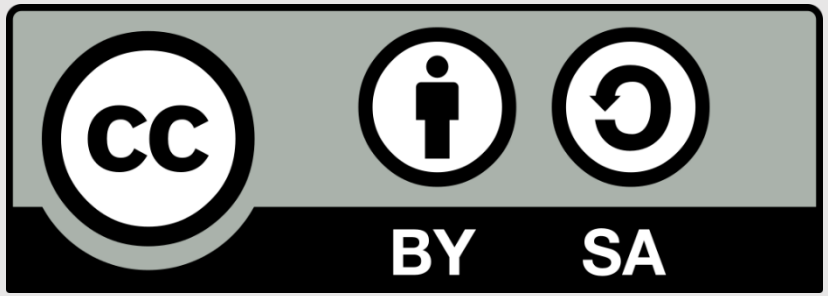
# UX in LLMs ;
# Data Discloser in Chatbots

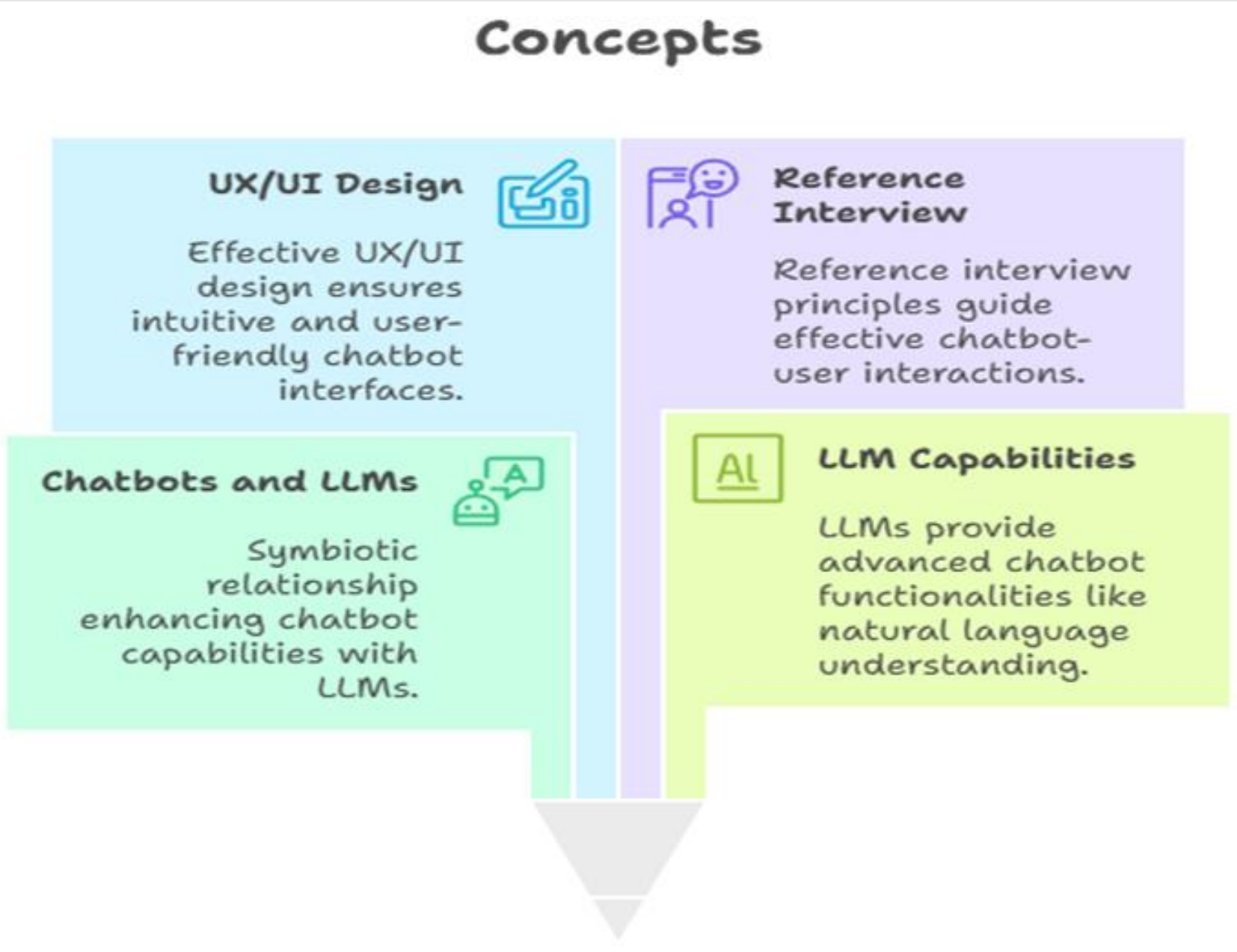
**Vahideh Zarea Gavgani (PhD)**
**Tabriz University of Medical Sciences**

**ORCID** https://orcid.org/0000-0002-3278-6859

**Email:** vgavgani@gmail.com, zarehv@tbzmed.ac.ir



# Concepts


UX/UI Design
Effective UX/UI design ensures intuitive and user-friendly chatbot interfaces.
Reference Interview
Reference interview principles guide effective chatbot-user interactions.
Chatbots and LLMs
Symbiotic relationship enhancing chatbot capabilities with LLMs.
AI
LLM Capabilities
LLMs provide advanced chatbot functionalities like natural language understanding.

# تعریف

- **چت‌بات‌ها** به عنوان "برنامه‌های رابط کامپیوتری هستند که می‌توانند مکالمه متنی یا صوتی با کاربران انسانی را حفظ کنند"
- به عنوان **عوامل توصیه‌گر** Recommending Agent
- **LLMها** بخشی از دسته وسیع‌تر هوش مصنوعی مولد هستند که به الگوریتم‌های یادگیری ماشینی اشاره دارد که می‌توانند از انواع مختلف محتوا، مانند متن، تصاویر و صدا، برای تولید محتوای جدید یاد بگیرند
- **UX/UI رابط کاربر و طراحی تجربه کاربر**
  - **مزایا و چالش ها**

# طراحی رابط کاربری UI و تجربه کاربری UX

- چت‌بات‌ها نوعی از رابط کاربری هستند که بجای کلیک روی دگمه و آیکون و یا پر کردن فرم از محاوره های صوتی و متنی استفاده می کنند که به آنها **(Conversational UI)**هم می گویند.
- در **UI( رابط کاربری)** چت‌بات می‌تواند به‌عنوان یک *مؤلفه تعاملی بین انسان و کامپیوتر* در وب‌سایتها، اپلیکیشنها یا سیستم ها استفاده شود مانند استفاده از آیکون، پنجره گفتگو، دکمه‌ها در طراحی ظاهری .
- در **UX ( تجربه کاربری)** باعث میشود محیط طوری طراحی شود که کاربر بتواند سریع‌تر، راحت‌تر و بدون نگرانی و استرس با چت‌بات مکالمه کرده و به هدفش برسد.

# نکات مهم طراحی UX/UI در چت باتها

- **شفافیت و وضوح (Clarity)** : برای اینکه سؤال‌ها و پاسخ‌ها کوتاه و ساده باشند.
- **مدیریت و کنترل مکالمه توسط کاربر:**طوریکه همیشه امکان انتخاب/لغو و خروج برای کاربر باسانی فراهم باشد.
- **امکان ثبت بازخورد (Feedback)** :وقتی پاسخ فراهم شد، نشست گفتگو به نتیجه رسید و کار انجام شد بتواند تأیید بدهد (مثلا حستجو برای یک موضوع به مقاله /مقالات مرتبط منجر شد)
- **برخورداری از طراحی تصویری /بصری جذاب** Visualized Design :استفاده ازعلایم، آیکونها، رنگهای مناسب، برای ارتقای ادراک و تعامل بهبود یافته. مثال :نگاهی به فلوچارت مکالمه کتابدار و هوش مصنوعی.
- **قابلیت شخصیت‌پردازی چت بات** ، ایجاد یک لحن دوستانه و داشتن حس هویت بات
- **قابلیت اتصال به پشتیبان انسانی** ، اگر کاربر نیاز به پشتیبان انسانی داشته باشد.

# مزایای UX

- **از بین رفتن حالت ارتباط استبدادی  و پیچیده بودن** ، Syntax MeSH
- **ایجاد محیط شخصی سازی شده برای محاوره بین انسان و چت بات ،**
- **ایجاد دسترسی پایدار،**

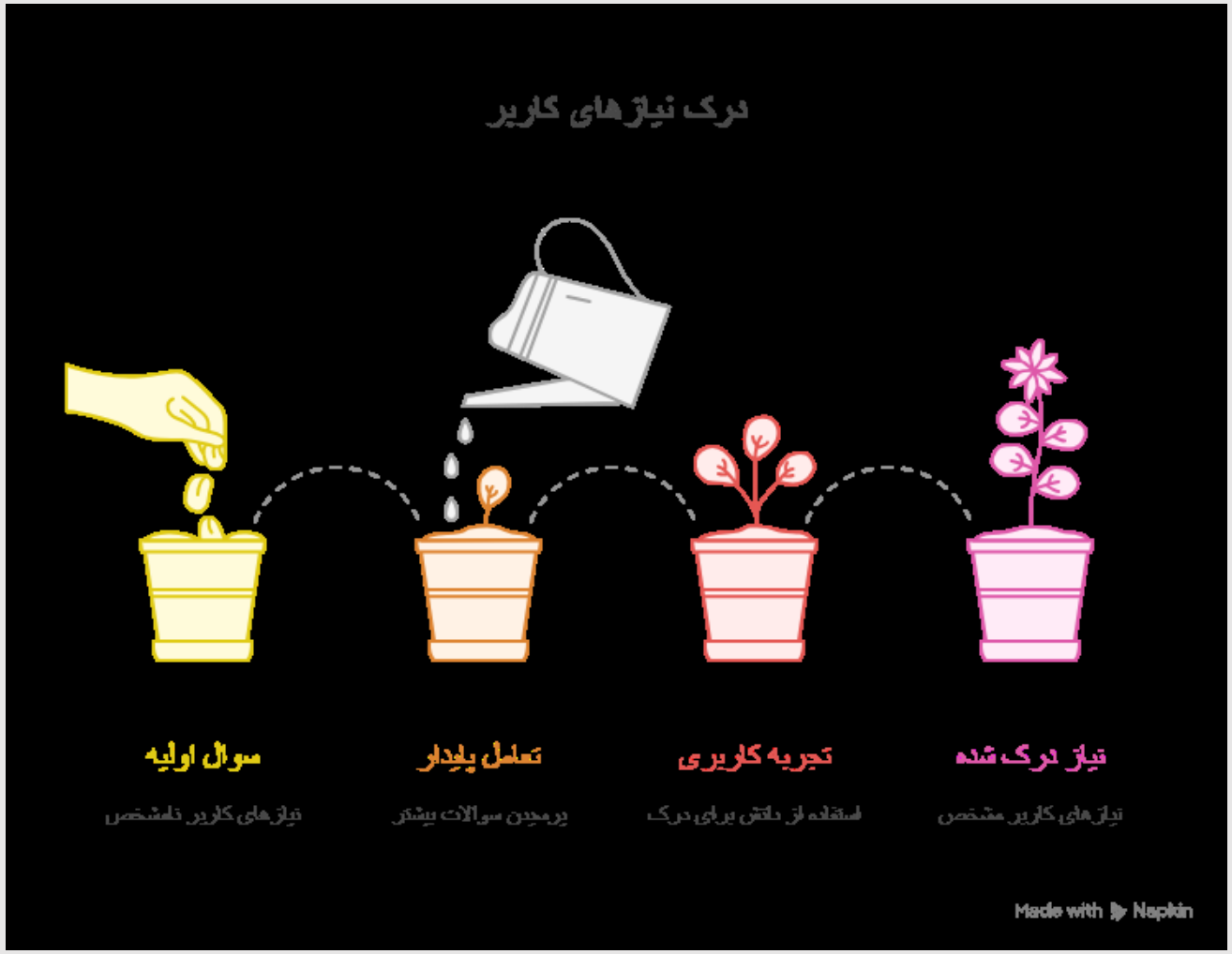
# چالشهای UX

- **احتمال وجود ابهام در درک چت بات از نیاز /سوال/ مکالمه**
- **ایجاد موازنه بین انسان و ماشین در طراحی HCI :**
- **احتمال ورود به حریم خصوص با استفاده از Nudges, Dark patterns**

# فلوچارت ارتباط در بحش مرجع یک کتابخانه با کتابدار مرجع

- 1-**شروع** -کتابدار آماده برای کمک و پاسخگویی با گشاده رویی ایستاده است و صمیمانه می پرسد می توانم کمک کنم؟ ،
- 2- مرحله سوال از کتابدار مرجع (خواستن سوال توسط کاربر)،
- 3-درک نیاز واقعی، تعیین سطح سوال ، زدون ابهام و بیان مسئله توسط کتابدار(تجربه ی کاربر)
- 4- بکارگیری منابع و ابزار ها، پیدا کردن و دسته بندی پاسخ(توسعه پایدار)،
- 5- ارائه پاسخ(تکامل)
- 6-گرفتن بازخورد( آیا همین را می خواستید؟ کافی است؟ اگر سوال دیگری باشد بپرسید)-
- 7- نقطه **پایان**
- ai

# فلوچارت هوش مصنوعی:




- مرحله ۱: بذر – (Seed) شروع تعامل
- مرحله ۲: جوانه – (Sprout) اولین تجربه کاربری
- مرحله ۳: درختچه – (Sapling) تعامل پایدار
- مرحله ۴: درخت بالغ – (Mature Tree) تجربه کامل و ارزش‌آفرین
- 6

- بخش دوم- ورود به حریم خصوص با استفاده از Nudges, Dark patterns

# User Privacy Violations in LLMs



**Rise of Data Disclosers**

Exposes private user information

**Unethical Nudges**

Manipulates user behavior unfairly

**Lack of Data Ethics**

Fails to treat users fairly

**Low Digital/AI Literacy**

Users can't navigate complexities

**Non-Adherence to Laws**

Rights are not adequately protected

# نگاهی گذرا به رفتار اطلاعاتی انسان با انسان و انسان با چت باتهای هوش مصنوعی در اطلاع یابی

- ارتباط با چت باتها شبیه به مصاحبه مرجع کتابدارن است. کتابدار برای درک نیاز واقعی کاربر مجبور است سوالاتی از وی بکند تا ابهام در سوال از بین برود و به شفافیت و دقت برسد.
- اما بندرت در طول تاریخ دیده شده(بجرات دیده نشده) کتابداری بخواهد به حریم خصوص و داده های کاربر نفوذ کند و از داده ها و پروژه وی استفاده غیر اخلاقی بکند.
- **اما در مورد چت باتها این نگرانی خیلی بزرگی است که شرکتهای سازنده چت بات مقاصد دیگری بغیر از دستیاری کاربر داشته باشند.**

# امنیت داده ها و حریم خصوصی کاربران

- **اخلاق داده**
  اخلاق داده به اصول درست و نادرست در جمع‌آوری، تحلیل، به‌اشتراک‌گذاری و استفاده از داده می‌پردازد.
- **اصول اساسی اخلاق داده عبارتند از :**
  - رضایت آگاهانه(Informed Consent)
  - شفافیت در کاربرد داده
  - پرهیز از تبعیض الگوریتمی
  - عدالت و انصاف در دسترسی به داده
  - احترام به حریم خصوصی حتی اگر قانون الزام نکرده باشد

- **افشای اطلاعات در مقابل حریم خصوص**
  فاش سازی اطلاعات و داده های کاربران نقض حریم خصوصی است، در یک سو کاربر قرار دارد که حقوقش نقض میشود در سوی دیگر شرکتها و چت باتها که به وسیله این حریم شکنی تقویت میشوند (یا برای اموزش مجدد چت بات یا برای فروش دادها)

# پارادوکس حریم خصوصی در LLMs

- **دسترس‌پذیر بودن ≠ رضایت استفاده**
- **امنیت دادهها و حریم خصوصی کاربران** یکی از موضوعات مهم دنیای دیجیتال **در بین کاربران، محققان و سیاست‌گذاران** است .
- شامل حوزه های گستره ای از داده های شخصی اجتماعی (هویت فردی) ، حسابهای مالی، بیماری و بالینی، امنیت ملی ، توسعه و تحقیقات ، تا تجارت و بازایابی، ...
- تنش بین تمایل یک شرکت برای به دست آوردن اطلاعات مشتری برای سفارشی‌سازی تجربیات و نیاز مصرف‌کننده به حریم خصوصی باهم تعارض و تضاد منافع یافته اند
- روش‌های روانشناسی و طراحی رابط کاربری (UI/UX) برای اشتباه انداختن کاربر درباره حریم خصوصی‌شان **"dark patterns"**یا تکنیک‌های **"nudge"**منفی

# اثر منفی فاش سازی اطلاعات بر تجربه کاربریUX

- چنگ و جیانگ : ریسک حریم خصوصی درک شده، سطح رضایت کاربران از چت‌بات‌ها را کاهش می‌دهد.
- رس و همکاران:"پاسخ‌دهندگان عموماً نگرانی‌هایی در مورد حریم خصوصی داشتند که بر فراوانی استفاده مورد نظر از چت‌بات‌ها تأثیر منفی می‌گذاشت."
- بعلاوه شناسایی کاربر بدون اطلاع یا رضایت وی یکی از رایج‌ترین و نگران‌کننده‌ترین مصداق های نقض حریم خصوصی در دنیای دیجیتال امروز است.

# معامله بین حریم خصوص و گرفتن اطلاعات مفید

- در ظاهر، اشتراک‌گذاری اطلاعات شخصی آنلاین ممکن است برای کاربران قابل قبول به نظر برسد. از دست دادن بخشی از حریم خصوصی در ازای شخصی‌سازی خدمات می‌تواند به عنوان یک تصمیم مصرف‌کننده سنجیده و حتی منطقی تفسیر شود.
- با این حال، حتی اگر کاربران از این بده‌بستان آگاه هم باشند، ممکن است در نهایت تصمیماتی برای افشای اطلاعاتی بگیرند که بعداً پشیمان شوند .

- کاربران همیشه از زمان و نحوه جمع‌آوری داده‌ها در طول تعاملات خود با یک شرکت و نحوه استفاده بعدی از این داده‌ها آگاه نیستند این واقعیت توجه دولت‌ها و نهادهای نظارتی را نیز به خود جلب کرده است

# قوانین کشورها

- General Data Protection Regulation (GDPR)
- California Consumer Privacy Act (CCPA),
- Brazil's LGPD,
- Japan's APPI
- **South Africa's Protection of Personal Information Act (POPIA)**
- **Canada's Personal Information Protection and Electronic Documents Act (PIPEDA)**
- **Turkey's Law on Personal Data Protection (LPDP)**
- اما ایران!

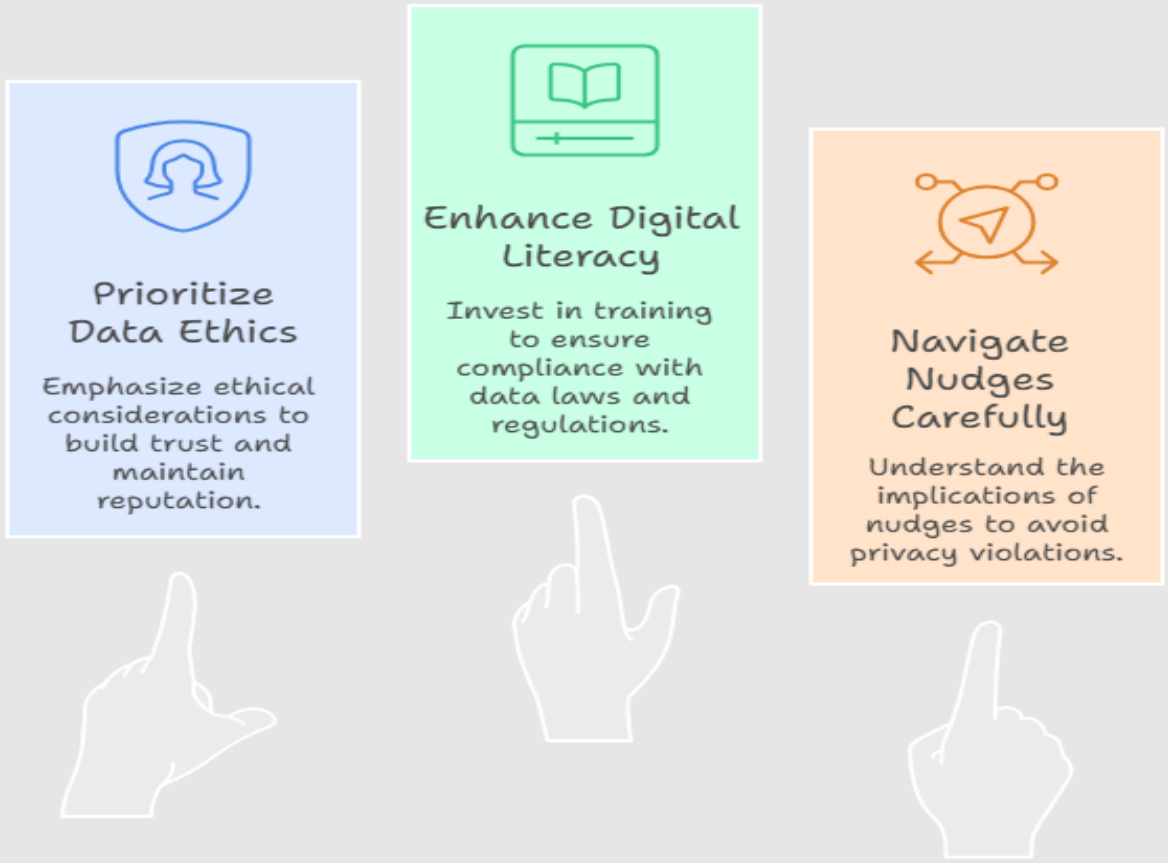
# آیا تصویب قانون کافی است؟

- **دستورالعمل‌های اخلاقی فعلی ارائه شده توسط دولت‌ها، در ارائه تاکتیک‌های عملی** که ثابت شده است کاربران را از رفتارهای افشای اطلاعاتشان آگاه می‌کند - و به طور بالقوه بر آنها تأثیر می‌گذارد - **شکست می‌خورند.**
- مسلماً، برخی از اطلاعیه‌های حریم خصوصی وجود دارند که به کاربران اطلاعاتی در مورد "نحوه و هدف جمع‌آوری، استفاده و مدیریت داده‌های آنها" ارائه می‌دهند. **با این حال، در واقعیت، کاربران به ندرت این اطلاعیه‌ها را می‌خوانند.**
- علاوه بر این، نشان داده شده است که **وقتی یک سیاست حفظ حریم خصوصی در یک وب‌سایت ارائه می‌شود، مصرف‌کنندگان ممکن است در نهایت اطلاعات شخصی بیشتری را افشا کنند.** دلیل این امر این است که مصرف‌کنندگان تمایل دارند به وب‌سایت‌هایی که یک اطلاعیه حفظ حریم خصوصی را نمایش می‌دهند، اعتماد بیش از حدی داشته باشند، زیرا معتقدند که از آنها بهتر محافظت می‌شود.

# سواد دیجیتال/هوش مصنوعی

- **بنابراین، نیاز به مداخلات سواد دیجیتال/هوش مصنوعی در بین کاربران و استفاده از ابزارهای ساده‌ای برای آگاه کردن کاربران از اطلاعاتی که قرار است به اشتراک بگذارند**، به ویژه با چت‌بات‌ها، وجود دارد.
- **چنین ابزارهایی می‌توانند رویه‌های اخلاقی شرکت‌ها را افزایش دهند و همزمان تصمیم‌گیری آگاهانه را در بین کاربرانی که ممکن است تصمیم بگیرند وسعت و عمق اطلاعاتی را که به صورت آنلاین به اشتراک می‌گذارند، کاهش دهند، ترویج دهند.**

# How should organizations approach data protection and ethics?



# Nudge

- **Nudge**یا **تلنگر رفتاری** یک مفهوم در اقتصاد رفتاری است که نخستین بار توسط Richard H. Thaler مطرح شد و ریجارد تالر آنرا ‘‘معماری انتخاب’’ نام نهاد. در حال حاضر این تکنیک در چت باتها بطور شایع در فرایند دادو ستد داده استفاده میشود.
- به زبان ساده :تلنگر رفتاری یعنی طراحی محیط تصمیم‌گیری به گونه‌ای که افراد را به سمت انتخاب مطلوب‌تر(از دیدگاه طراح یا کاربر) سوق دهد، بدون اینکه آزادی انتخاب آنها محدود شود.

- **تلنگر های رفتاری در چت باتها مثبت و منفی هستند.**
- تلنگر های مثبت شامل مواردی مانند :
  - چت باتهای سلامت محور و مرتبط با سبک زندگی
  - چت باتهای آموزش و یادگیری محور

- **تلنگرهای منفی غالبا برای فریب کاربر طراحی شده اند**

# تکنیک‌های مثبت Nudge در چت‌بات‌ها

- **پیش‌فرض‌ها (Defaults) :**تنظیم گزینه‌ای به عنوان پیش‌فرض (مثلا: بله، می‌خواهم یادآوری روزانه دریافت کنم).
- **قاب‌بندی (Framing) :**تغییر بیان جمله برای اثرگذاری بیشتر (۹۰٪ افراد همین مسیر را انتخاب کرده اند).
- **یادآوری و زمان‌بندی (Reminders) :**ارسال پیام در زمان مناسب برای افزایش شانس اقدام.
- **اجتماعی‌سازی** محرک اثبات اجتماعی **: (Social proof)**نشان دادن اینکه دیگران چه انتخابی کرده‌اند.
- **خرد کردن کارها (Chunking) :** پیشنهاد قدم‌های کوچک به جای هدف بزرگ مثل بیا گام بگام جلو برویم.

# استفاده‌های منفی از Nudges در چت‌بات‌ها

- **تلنگرهای فریبنده(Deceptive Nudges)**
- طراحی پیام یا گزینه‌ها به‌گونه‌ای که کاربر بدون دقت یا آگاهی تصمیمی بگیرد که به نفع سیستم یا شرکت است، نه خودش.
- مثال: وقتی چت‌بات می‌پرسد:برای تجربه بهتر، اجازه دسترسی به مخاطبانت را می‌دهی؟ جواب مثبت به این سوال را هایلایت و جذاب می کند که کاربر آن را گزینه ترجیحی تشخیص می دهد

- **استفاده از پیش‌فرض‌های خطرناک(Dark Defaults)**
  - انتخاب پیش‌فرض را طوری تنظیم می‌کنند که کاربر ناخواسته داده‌های شخصی‌اش را به اشتراک بگذارد.
  - مثال: در ثبت‌نام، گزینه‌ی "*اشتراک‌گذاری داده‌های سلامت با شرکای تجاری* "به‌صورت پیش‌فرض فعال است و کاربر باید (opt-out)کند یعنی عمداً غیرفعال کند اما بی توجیهی می کند.
- **قاب‌بندی دستگاری شده برای گمرا ه کردن(Manipulative Framing)**
  - اطلاعات را طوری قاب‌بندی می‌کنند که کاربر حس کند اگر اجازه دسترسی یا خرید ندهد، چیز مهمی را از دست می‌دهد.
  - مثال: ۹۰٪ کاربران ما با به‌اشتراک‌گذاری داده‌هایشان تجربه خیلی بهتری داشته‌اند
  - موقعی داری در یک مکالمه به جای حساس می رسی ازت تقاضای لاگ-این می کندو آنقدر طول میدهد مجبور شوی تسلیم شوی و با ایمیل واقعی لاگین شوی.
- **تایید اجتماعی ساختگی(Fake Social Proof)**
  - چت‌بات ممکن است وانمود کند «همه این کار را کرده‌اند» تا فشار اجتماعی به کاربر وارد شود.
  - مثال: «دوستانت همین الان این سرویس را فعال کردند!» (درحالی‌که واقعیت ندارد).
- **بازی با زمان و خستگی تصمیم‌گیری(Decision Fatigue)**
  - چت‌بات عمداً مسیر رد کردن یا لغو اشتراک را طولانی و خسته‌کننده طراحی می‌کند تا کاربر تسلیم شود.
- این همان چیزی است که به آن **Dark Patterns**در طراحی رابط می‌گویند.

# پیامدهای منفی این نوعNudges

- **نشت اطلاعات حساس** :داده‌هایی مثل موقعیت مکانی، سلامت، یا عادات فردی ناخواسته به شرکت‌ها یا اشخاص ثالث می‌رسد.
- **از بین رفتن اعتماد** :وقتی کاربر بفهمد فریب خورده، اعتمادش به کل سیستم از بین می‌رود.
- **آسیب به حقوق کاربر** :نقض اصل رضایت آگاهانه Informed Consent
- **مشکلات حقوقی و اخلاقی** :نقض قوانی حفاظت دادههای شخصی کشورها مانند GDPRدر اروپا

◦پایان قسمت اول نشست

# پایان