

Respondent Jan Misák, datum 13.10. 2023

Jaké nástroje nebo prvky AI využíváte v rozhlase a v rámci aplikace mujRozhlas?

Ted' jsme ve fázi mezi starou verzí doporučováním, kterou ted' máme od externího dodavatele. Je to česká firma Recombee, která má základy na ČVUT. Jde o hodně primitivní verzi, která nám už běží dlouho. My jsme to brali spíš jako *proof of concept*, testování v produkčním prostředí, jak to bude vypadat, jestli o to bude zájem, jak to bude drahé a podobně. Ted' jsme zrovna ve stavu, kdy vypršela lhůta pro výběrové řízení. Máme sepsané body a záměr, který chceme od dodavatele, který výběrové řízení vyhraje, dodat. Dodavatel, který výběrové řízení vyhraje, má tuším 3 měsíce na to, aby implementoval to, co chceme, tedy novou verzi doporučováním. My to mu říkáme doporučováním 2.0.

Jak fungovalo doporučováním doted' a jaké jsou ty nové požadavky?

To, co máme ted' od Recombee pro mujRozhlas, je doporučováním souvisejících audií k jednomu každému audiou. Používáme to asi na dvou místech v rámci mujRozhlasu, což je audioportál Českého rozhlasu, a to v aplikaci i na webu. „Ke každému jednomu audiou máme nasazený trackovací skript, takže odesíláme anonymně data našemu dodavateli o tom, co lidi, kteří poslouchali audio, poslouchají za další audio. Ten dodavatel sbírá data na svém cloudu a dodává nám to už jako hotovou službu. Funguje to tak, že mi mu datovou vrstvou odešleme jedno naše audio a on nám k němu pošle zpátky 10 audií, které máme zobrazit na webu v rámci toho doporučováním.“

A to už teda funguje...

To už funguje. Zkoušeli jsme postupně asi 3 různé algoritmy. Recombee má v cloudu svoji službu, kde jim běží jejich výpočetní výkon. Mají tam nastavené algoritmy, které průběžně vylepšují, neustále indexují nové články, vztahy mezi nimi, indexují všechny data a metadata jako titulek, popisek, článek, nějaké datum vydání, pořad, ze kterého to je atd. Pak indexují zájem o ty audia, kdo co poslouchá v souvislosti, jako klastr těch audií. To znamená, že když poslouchám A, tak většinou k tomu poslouchám B. Všechny tyhle data indexují a na základě toho na dotaz „dej nám doporučení k tomuto audiou“, nám umí doporučit dalších 10 souvisejících. K tomu máme ještě vlastní „nadstavby“ a vlastní pravidla. Pokud máme audioseriál, který má 10 dílů, tak tam máme vlastní logiku ve stylu, že nebereme jenom dalších 10 souvisejících audií, ale že máme nastaveno, že chceme jako to doporučení podávat nejbližší další epizody tohoto seriálu. Takhle tam máme pár těch vlastních logik.

A když vám to oni pošlou, co s tím potom děláte?

Zobrazujeme to myslím na dvou místech. První je na detailu audia na mujRozhlasu. Když se najede na jakýkoliv detail audia, dole je box, který se jmenuje *Ostatní také poslouchají* a to je typicky to doporučení. Tohle je těch 8 až 10 doporučených audií. To je první místo, kde je zobrazujeme a druhým místem je fronta, která je hodně podobná Spotify. Když začneš poslouchat audio, tak my máme implementovanou nějakou frontu jako v plovoucím přehrávači, a když máš nastavené automatické přehrávání potom co doposloucháš to audio, tak my na tebe umíme hrnout další a další související audia. Takže ta myšlenka je, že člověk zkonsumuje samozřejmě víc audií. Doposlechne jedno a ten rekomendační script si myslí, že těch dalších 10 toho uživatele bude zajímat, takže to máme nastavené tak, že hrají automaticky hned potom. Když si audio člověk nevypne před koncem, tak mu začne hrát to první nejrelevantnější doporučení. Existují i data úspěšnosti tohoto doporučení. Dost nás překvapilo, že to má dobrou úspěšnost. Výrazná většina lidí, kterým se začne přehrávat první

doporučené audio, tak ho nestopne, ale pokračuje v poslechu. Třeba zhruba 60 % z těch lidí dojde u toho prvního doporučeného audia třeba až do tří čtvrtin.

Když přijdeš na tuhle stránku, tak si tu funkci musíš automaticky zapnout, nebo vypnout?

Defaultní chování je, že to je zapnuté. Ty, když přijdeš úplně poprvé na mujRozhlas, pustíš si audio a odejdeš od počítače, tak máš puštěné automatické přehrávání a kdybys odešla na 8 hodin, tak ti přehraje všech těch 10 doporučených audií. V mobilní aplikaci je to stejně jako na webu. Za hodně důležité považujeme tu zmíněnou frontu. Ten box se zobrazováním *Ostatní také poslouchají* je sice fajn, ale je to takový doplněk, který je dole a ke kterému dojde méně lidí. Primární účel a využití doporučování je mít to v té frontě. Stejně jako na Spotify. Většina trafficu Spotify, většina přehraných písniček na Spotify nejde z toho, že si někdo dá *play* a poslechne si jednu písničku a pak si tím *play* pustí jednu za druhou, ale většinou taky spadne do té fronty, která je tam naservírovaná. To byla, je a taky bude naše idea. Naučit se, co ten člověk poslouchá, co poslouchá pravidelně a podávat mu to do fronty tak, aby už nemusel nic hledat na webu, jen by si pustil ty svoje audia a oni by mu tam padaly automaticky. Pak máme v tom doporučování ještě další vlastní pravidla, například aby tam nepadaly expirovaná audia, co nejdou přehrát, protože nějaká audia například z Vltavy jsou omezená autorskými právy, takže jsou dostupné třeba jenom 30 dní a pak musí odevšud zmizet, abychom nenabízeli něco, co bude házet chybu. Anebo tam filtrujeme to, co už má uživatel doposlouchané nebo rozposlouchané a dál mu nabízíme vlastně jenom nová audia, aby je mohl objevovat. Nevracíme mu tam to, co už má doposlechnuté do půlky a chtěli bychom, aby si to doposlechnul, protože na to má na webu i v aplikaci jiná místa. Ty dříve poslouchané si může dohledat v historii.

Co teď plánujete dál?

Plánujeme personalizaci. Protože dnes to nemáme buď vůbec, nebo velmi jednoduše personalizované. To znamená, že každý uživatel vidí většinou to stejné, co vidí jiní uživatelé. Personalizace spočívá jenom v tom, že mu tam filtrujeme ta jeho doposlechnutá audia. Takže v tom by byla ta personalizace, že když doposlouchá těch deset doporučených audií, tak dostane deset jiných než člověk vedle něj. Takže ta personalizace tam je, ale pořád nedokonalá.

Jo, takže to je ale stále z těch stejných klastrů?

Ano. Teď už máme uživatelský účet na mujRozhlase, takže ideální je tlačít lidi, aby byli přihlášení a my jsme měli k nim data a nikdy je neztratili, abychom jim mohli relevantně doporučovat. Takže to je na základě uživatelského účtu a pro ty nepřihlášené to bude na základě nějakého ID z cookies v prohlížeči nebo z ID z mobilní appky, kdy si budeme pamatovat, že to je uživatel s XYZ s tímhle ID a taky mu to budeme co nejlíp personalizovat podle historie doposlouchaných audií. Jenom by to nebylo dokonalé, protože když si resetuje nastavení v prohlížeči nebo změní počítač, tak už ho vidíme jako nového uživatele. Takže tolik k personalizaci. A pak, chceme mít víc logik. Dnes máme jenom dvě. Prostě to doporučování má aktuálně jednu logiku jako jeden ten algoritmus ke všem audiím. Ať už je to zpravodajské audio, divadelní hra nebo zprávy Zelené vlny, cokoliv, tak podle stejného algoritmu tam padají ta audia. Což není dobře. Protože u zpráv chceme posluchači podávat další doporučená audia, která mají třeba větší důraz v tom algoritmu na datum. Například když poslouchá zprávu, že se na Pražském hradě propadl strop, aby mu to nedoporučilo deset let starou zprávu se stejným obsahem, ale spíš aby mu to doporučilo nějaký širší tematický

a relevantní okruh. U těch zpráv jde člověku o aktuálnost, tam ten obsah zastarává. To teď nemáme aktuálně. Máme tam jenom nějakou jednoduchou pseudologiku pro dětský obsah, kdy tagujeme pohádky jako obsah pro děti a k obsahu pro děti doporučujeme zase jenom další obsah z klastrů *série pro děti*. Takže dva klastry tam máme aktuálně a nově poptáváme minimálně čtyři klastry na to doporučování, takže čtyři upravené algoritmy. První zpravodajský. To jsou typově zprávy nebo aktuální publicistika, třeba 20 minut Radiožurnálu, nějaké aktuální zpravodajské rozhovory a podobně. S důrazem na zpravodajský obsah, a nikoliv třeba dramatická díla, a čas. Pak druhá logika, tomu říkám literárně-dramatická, to jsou typově jako Vltava, Wave, prostě volnočasové podcasty, nějaká díla, dramatizované divadelní hry, prostě umění. Pak dětská, to bude podobné tomu, co máme teď a pak třetí publicistická, to jsou hodně volnočasové jako třeba rozhovory z Dvojky jako Blízká setkání nebo rozhovory na Wavu. Máme požadavek na vítěze toho doporučování, že v rámci úvodní implementace musí s námi udělat workshop a navnímat naše potřeby, udělat nám a uzpůsobit ten doporučovací mechanismus na míru, takže je možné, že nám třeba doporučí, že máme mít ne čtyři, ale šest a najde nám ještě dva další segmenty. Každopádně je to pokrok vůči aktuálnímu stavu, kde máme jenom dva.

A ten uživatel dostane vlastně kombinaci všech? Pokud není dítě...

Podle toho, co poslouchá. To není *per user*, ale *per to audio*. To znamená, že když začneš poslouchat Blízká setkání, tak ti algoritmus pro volnočasové doporučování vybere zase volnočasová audia z Blízkých setkání. Když si ráno poslechněš zprávy, tak ti to doporučí další zpravodajský obsah. Nejde to cílit *per user*, protože to, že si ráno poslechněš Blízká setkání neznamená, že chceš celý den, kdy posloucháš zprávy, k tomu dostávat nějaké rozhovory. Musí to být vždy podle aktuální situace. A pak tady máme nějaké extra nadstavby, třeba na ty díly seriálu. Máme vymyšlená nějaká další editorská zvýhodňování. Typicky když na Wavu vyjde nový podcast, tak ho naboostujeme. Nastavíme, že pro publicistickou logiku chceme naboostovat obsah, který spadá do téhle série podcastů, aby se vracel častěji třeba, protože to je novinka a chceme, aby lidí do toho víc zabředli. Takže to je nějaké editorské zvýhodňování. A to je to, co chceme. Teď boostujeme pouze ty díly seriálů, když si poslechněš jeden díl, tak dostaneš automaticky těch deset doporučených. Pak jsme chtěly blacklistování vybraných audio pořadů. Chtěli jsme mít v administraci taky možnost něco blacklistovat, aby se to nedoporučovalo. Dále požadujeme možnost propojení mezi administračním systémem toho doporučování a našim administračním systémem na obsah Drupal. To znamená, že by to bylo provázané s editorem a jeho prací a jakmile editor dá ručně do boxu na homepage „Editor doporučuje těchto šest audií“, tak se o něco zvýší pravděpodobnost, že se bude těch šest audií vracet v doporučení. Protože víme, že to je teď aktuálně to nejlepší, co máme tento týden, ten top obsah a chceme to *naboostovat*, aby to slyšelo co nejvíc lidí.

Tohle teda bude platit pro mujRozhlas a propíše se to i na iRozhlas?

Tohle se týká jenom audií. Ta zakázka na doporučovací mechanismus je sice společná pro iRozhlas a mujRozhlas, ale oni to budou vlastně dvě nezávislá doporučení, protože my doporučujeme k audio audio, kdežto iRozhlas chce ke článku články. A je to úplně jiný obsah. iRozhlas má třeba jenom články bez audia. Oni mají jenom články, takže to budou dvě nezávislá doporučení. Další požadavek je na prostory, kam to chceme zobrazovat. Říkal jsem, že to máme hlavně na dvou místech, na detailu audia a ve frontě. Máme a nově chceme mít: *stránka detailů audia, poslechněte si také, fronta audií*, to máme teď, pak chceme mít *push notifikace pro web a aplikace*, typicky „nezmeškejte tohle“, to znamená

nějaké konkrétní doporučení audia, které trenduje a je to personalizované tak, že by to bylo audio, které uživatel ještě neslyšel. Chceme mít automatizaci v rámci doporučování, že třeba jednou za dva nebo tři dny vyskočí uživateli personalizovaná *push notifikace* v mobilní aplikaci, která mu doporučí třeba jedno z deseti audií, které ještě neslyšel a je to nějaké top audio, o kterém víme, že ho uživatelé poslouchají. Takže doporučení jako *notifikace o trendujícím audiu*, bych to nazval.

Pak na tematické stránky. To máme na mujRozhlase třídění obsahu podle témat, jsou tam například *zprávy, příroda, věda a technika* a podobně. Tak tam chceme přidat box *ostatní poslouchají* a bylo by to doporučení audií k aktuálnímu tématu na základě pořadů, které vysílají aktuální téma a zase by to bylo to, co člověk ještě neslyšel. Prostě výběr nejrelevantnějších audií pro to téma. Že když uživatel přijde na téma *věda a technika*, tak tam nebude mít nahoře chronologický výpis nejnovějších audií z tématu *věda a technika*, ale ten nejnovější výpis bude: „Vybrali jsme pro vás z tohoto tématu, protože si myslíme, že vás to s největší pravděpodobností zajímá“. Takže to je další místo, kde to chceme ukazovat.

Pak tady máme pořady, *stránka s detailem pořadu*. To není doporučení audia, ale tady chceme doporučovat pořad. Doteď jsme se bavili jenom o doporučování audií k audiu a teď chceme mít i doporučování pořadů k jednomu pořadu. Například k Hostu Lucie Výborné - Blízká setkání. Bude to vybírat, které pořady jsou si nejpodobnější podle nějakého algoritmu. Buď třeba obsahově, když je tam podobný tematický záběr, nebo to bude komparativní sledování, že „uživatelé, kteří poslouchají pořad A, tak nejčastěji poslouchají další audia z pořadu B, takže my vám doporučujeme tento pořad B.“ Prostě nejpodobnější pořady podle poslechovosti nebo něčeho.

Další jsou *odebírané pořady*. Dneska uživatel může odebírat pořady. To znamená, že mu pak padají nejnovější audia z těch pořadů do nějaké sekce na homepage a my mu chceme doporučovat další. Myslím si, že to má třeba Youtube, že když odebíráš 10 kanálů na Youtube, tak tam někde máš prostor *odebírejte taky tyto*. Ono to k jemu odebírané množině kanálů vybírá další, které jsou podobné a mohly by ho taky zajímat na odběr. Takže to je další prostor, kde chceme využít ten doporučovací mechanismus. To stejné je u seriálů. K seriálu chceme doporučit další seriály. Nejenom ve smyslu, že k dílu seriálu chceme doporučit další díly. Ale například třeba Hamlet od Shakespeara má 10 epizod, takže k Hamletovi bychom doporučili třeba Othella.

Tohle je základ, který máme nadefinovaný pro ten start. Když to výběrové řízení teď někdo vyhraje, tak do tří měsíců bychom ideálně chtěli tohle všechno od nich dodat. A neznamená to, že to je jako jediné. To výběrko je na 3 roky. Oni nám implementují tohle, pak to 3 roky poskytují, ale my tam máme ještě peníze na nějaký rozvoj. Takže když dojdeme k tomu, že chceme něco přidat, například nějaký nový box nebo funkce, tak si je můžeme nechat ještě implementovat. Tohle je to minimum, které určitě musí být.

Pak v rámci toho doporučování chceme nějaké žebříčky, že bychom to zobrazovali jako nějaké boxy, třeba na homepage. Takže někde máme třeba *trending* na homepagi. Jakože „ostatní poslouchají a tys to ještě neslyšel“. Potom oblíbené pořady ostatních posluchačů, takže to je doporučování těch pořadů k nějakému aktuálnímu trendu. *Speciálně pro vás*, to jsou audia, co uživatel neslyšel a co jsou podobné tomu, co má ten člověk v historii. Takže tam vybíráme *pro vás*, a to může být třeba box na homepage, co se mění třeba každý den. *Nejposlouchanější za posledních 5 dní*, to je vlastně jednoduché, to může být i podle Google Analytics, prostě nejposlouchanější jenom za 5 dní a okrouhané zase o to, co už slyšel. Chceme tam mít nějaký *weekly catch up reminder* ve smyslu, že „tohle audio trendovalo minulý týden, bylo to top a vy jste ho neslyšel, tak si myslíme, že byste ho mohl slyšet,

protože si myslíme, že vás zajímá“. Nebo nějaké *doporučování oblíbené kategorie*, že pokud víme, že uživatel poslouchá z kategorie *věda a technika*, tak bychom mu na homepage přidali box *věda a technika* a dávali mu tam audia z jeho oblíbené kategorie, nějaká prostě top doporučení, co by ho mohla zajímat.

Potom tady máme nějakou indexaci dat, tedy „co všechno chceme za data posílat tomu dodavateli, aby on ho indexoval a podle toho vrátil ta doporučení“. Tak tady je *historie* toho uživatele, *rozposlouchanost*, to znamená, když doposlouchal audio třeba do čtvrtiny. Potom *přepisy textů audií*, jednak přepisy článků, které k tomu audio vznikly. Oni je naindexují a podle toho textu hledají podobné. Ale i přepisy přímo těch audií, že je přepíšeme přepisovačem do textu a oni je naindexují a podle toho budou hledat podobné audio, aby to taky zohlednili v tom doporučování. *Odebírané pořady uživatele*, *oblíbená audia uživatele*, *oblíbené stanice uživatele*, co třeba poslouchá živě nebo z jakých stanic nejčastěji poslouchá *audio on demand*. To je asi už všechno, co tady máme jako ten první výkop k mujRozhlasu, co chceme mít co nejdřív.

A ta audia jsou tedy ze všech stanic, to se nijak nerozděluje?

Na mujRozhlase, tím že to je poslechový portál Českého rozhlasu, tak my ta audia umíme třídit podle stanice, ale default je, že to má být třeba jako Netflix nebo Spotify, že prostě přijdeš a sleduješ to audio a vybíráš si podle toho, co je to za audio, a ne podle toho z jaké je stanice. Jako dá se to tam vidět, z jaké stanice každé audio je, ale není to primární účel, podle kterého by to ten uživatel měl poslouchat. Účel, se kterým jsme kdysi dělali mujRozhlas byl, že kdysi byl jenom webový prostor jednotlivých stanic, takže uživatel, když chtěl poslouchat audia, tak si musel vybrat stanici, jít na její web a hledat audia. A toho jsme se chtěli zbavit. Chtěli jsme jeden prostor pro všechny posluchače audia nezávisle na tom, co mají rádi za stanici. Z toho vlastně pramení gró toho, čím je mujRozhlas dnes.

Co podle tebe využívání těchto prvků přinese posluchačům? Jaké výhody nebo jaké etické výzvy?

Výhody, že jim ušetří čas a že je zbaví obsahové paralýzy. To je dneska největší problém třeba Netflixu. Večer si zapneš Netflix a půl hodiny brouzdáš, co by sis pustil, a to je vlastně to, čeho se chceme zbavit. My chceme, aby ten uživatel přišel a v těch několika boxech, které uvidí během pár vteřin, co proskenuje tu homepage nebo nějakou svojí uživatelskou sekci, tak aby viděl, co všechno teď relevantního Rozhlas nabízí. My jsme mu předžvýkali to, co ho s největší pravděpodobností bude zajímat a on si z toho vybral. Takže aby mu to ušetřilo čas a rozhodovací paralýzu. Pak se samozřejmě chceme vyvarovat toho, aby ho to uzavřelo do nějaké sociální bubliny, aby mu to nepodávalo jenom obsah, který on vlastně chce a blokovalo ho to od prohlížení obsahů, které třeba obsahují třeba opačný názor. Ten doporučovací algoritmus nesmí uživatele uzavírat v té jedné bublině. Například, aby mu třeba, zobrazoval rozhovory jenom s lidmi, se kterými názorově souzní, ale aby mu ukazoval rozhovory z celého politického spektra a podobně. Takže to je ta výzva pro nás, neuzavírat uživatele do bubliny.

Uživateli to přinese i větší rozhled. Dneska ten uživatel je v záplavě obsahu a díky tomu mu spousta věcí unikne. On třeba neumí rozpoznat na mujRozhlasu nebo i iRozhlasu, co je to důležité. Třeba ho úplně mine nějaká problematika během toho týdne a my stejně chceme, aby se o ní dozvěděl. Takže když vidíme, že ho mine nějaká důležitá věc, která ten týden rezonovala, tak chceme, aby on o ní stejně zavádl tím, že mu jí pošleme, nebo že mu jí dáme na homepage do nějakého boxu, aby nepřišel o nic aktuálního, co by měl vědět jenom proto,

že třeba nějakou dobu nebude konzumovat informace. Chceme, aby ta důležitá informace nezapadla, ale uměla si ho najít. Když poznáme, že je relevantní.

Máte tam nějaký mechanismus, který by tohle vyřešil?

To by byly třeba ty *push notifikace*. Třeba každé dva, tři dny to vybere jedno audio, co uživatel ještě neslyšel, z těch co nejméně rezonovala a pošle mu to „tohle bylo trendy, ty jsi to neslyšel, asi by bylo fajn to slyšet, protože ostatní to považují za důležité“.

A pošle mu nejen ta audia, která by nemusel považovat za nejdůležitější z toho týdne, ale třeba i ta, která jsou názorově opačná?

Třeba taky.

Takže ty notifikace řeší obojí?

Přesně tak. Notifikace neřeší, jestli ten, kdo byl v tom audio, jestli je levičák nebo pravičák, ale že to byl prostě nějaký rozhovor, který rezonoval třeba proto, že levičáci ho nesnášeli a pravičáci ho adorovali. A nás nezajímá, kdo je ten člověk, kterému to přijde, ale on prostě neslyšel téma, které bylo důležité, takže my mu ho dodáme. Ještě jsem vlastně nezmínil newsletter, protože to taky souvisí s tím, co říkáš. My každý pátek posíláme newsletter uživatelům. Jmenuje se to Tipy k poslechu a vybírá to dneska člověk, Anička Vošálíková a my bychom chtěli, aby nám s tím pomohlo to doporučení, aby část toho newsletteru byla stále vybíraná člověkem, ale část toho doporučeného obsahu by byla vybrána strojově z toho rekomenačního enginu. Takže by tam klidně mohlo být zase jako „týden jsi neposlouchal audio, týden jsi nebyla na rozhlase, tak tady máš ten pravidelný páteční newsletter a tady máš výběr pěti audií, co by ti za ten týden neměly zapadnout.“

S tím souvisí i moje další otázka. Co podle tebe využívání těchto prvků přinese pracovníkům rozhlasu. Jaké výhody a etický výzvy?

Já doufám, že pomoc, protože tady v rozhlase máme problém, že máme málo lidí a máme malý budget, nestíháme pracovat a dělat všechno to, co bychom chtěli, takže já doufám, že to pomůže těm lidem, kteří jsou zahlceni prací a že jim to ulehčí část práce, kterou to udělá za ně. Ti lidé se budou moct dál věnovat něčemu, co za ně ta rekomendace neudělá. Prostě něco, co ti zaměstnanci dneska nestíhají dělat kvůli tomu, že musí vybírat *top nejlepší audia*. To už nebudou muset dělat, uvolní se jim ruce a budou moct dělat něco jiného důležitého. Samozřejmě, vždycky tu budou jako nějací odboráři, co budou řvát a bát se, že to bude připravovat lidi o místa, ale já si to nemyslím. My tady máme tak málo lidí, že jakmile by se jednomu člověku uvolnily ruce a měl čas dělat něco jiného, tak my tady pro něj máme miliardu další práce, co budeme potřebovat, aby dělal místo toho.

Tak to byly ty výhody a teď ty etické výzvy.

Tak asi jenom to zavírání toho člověka do bubliny. Nebo třeba tomu říkám „Vltava model“. Na Vltavě vždycky ti autoři říkají: „My tady děláme strašně drahý a důležitý obsah, službu kultuře a společnosti, jenže to nikoho nezajímá. Nemůžete měřit audio jenom podle poslechovosti, protože to, že my tady vyrobíme něco, co nikoho nezajímá, neznamená, že to není dobrý, ale naopak to je třeba zachování kultury pro další generace, že to bude doceněné až za 10 let a je hloupost, že vy nám říkáte, ať toho teď děláme míň jenom proto, že to lidi aktuálně neposlouchají.“ Nejenom obsah, který ten doporučovací mechanismus vyhodnotí jako relevantní, tak je jediný relevantní obsah. Může tady být další obsah, který se z nějakého důvodu nebo omylem nebo náhodou nedostane do té smyčky, že by prostě začal *boostovat*,

a díky tomu ho ten doporučovací mechanismus bude ještě více doporučovat, ale musíme prostě sledovat a umět mezi lidi šířit i další obsah, který není prvoplánově dobrý v tom smyslu, že ho lidi začnou okamžitě sdílet a rezonuje, ale je prostě kvalitní a taky se potřebuje dostat mezi ty lidi. Takže nesmíme podléhat prvoplánovitosti a šířit jenom jako nějaké pozlátko a clickbait, který bude rezonovat. Musíme šířit i ty kvalitní věci, které se třeba šíří pomaleji ze začátku, ale dlouhodobě o ně je zájem, protože jsou to třeba ty kvalitní věci z Vltavy, co si zaslouží být zachovány a dál natáčeny, protože chrání kulturu České republiky.

Jsou podle tebe novináři a editoři přepravení pracovat s prvky nebo s nástroji AI?

Já myslím, že ano, že to je téma, které u nás hodně rezonuje a když se o tom bavíme s lidmi z baráku, tak mi přijde, že jsou erudovaní. Že znají i AI nástroje jako třeba Midjourney nebo ChatGPT, hodně z nich si to i zkoušelo ve free verzi, hrát si s tím, dívali se, jak to funguje, byli třeba na nějakém interním workshopu od kolegy Honzy Cibulky, kde jim je ukazoval. Takže já myslím, že lidi z rozhlasu a novináři, zaměstnanci médií jsou erudovanější, než průměrný člověk venku na ulici. Když srovnávám lidi v rozhlasu a třeba kamarády, tak kamarádi vůbec neřeší AI, nezajímá je to, skoro nikdo z nich si nezkoušel ChatGPT, ale tady v baráku mi přijde, že skoro všichni. Myslím, že zájem o to je a hrozně se mi líbí, že novináři jsou zvědaví, zkouší to a hledají, jak by jim to mohlo pomoci.

Budete je nějak i školit?

Už je školíme teď, třeba jsme dělali Midjourney a ChatGPT na iRozhlase. Tam jsou třeba vyškolení kolegové na to, jak si můžou nechávat navrhovat posty na sociální sítě od ChatGPT. Mají tam třeba sbírku promptů, kdy zadají prompt a celý článek, a navrhne jim to třeba pět variant nejlepšího postu na Facebook. Takže tam třeba experimentují s tímhle.

A ohledně toho mujRozhlasu, až budete zavádět ty doporučovací algoritmy, tak je taky v plánu nějaké školení pro tvůrce toho obsahu nebo pro editory?

Já si myslím, že by to asi ani nepotřebovali... nebo jako k čemu by tvůrci podcastu na Wavu bylo školení o algoritmu, podle jakého se mu na webu mujRozhlas doporučuje audio z dalšího? Možná jsem ve své bublině a přijde mi na první dobrou, že to není potřeba a možná naopak za mnou někdo přijde, že to chtějí vědět, že je to zajímavá, protože chtějí vědět, proč se u jejich podcastu zobrazuje podcast B a ne podcast C jako doporučení. Možná je to jenom moje slabina, že nedohlídnu takhle daleko, že by o to lidi měli zájem a k něčemu by jim to třeba bylo.

A těm editorům?

Tam je určitě budeme školit k tomu, co potřebují. Jakmile bychom implementovali mechanismus toho propojení ve smyslu, že co si dá editor na homepage své stanice, třeba top audia za tento týden na stanici Radiožurnál, tak jim musíme vysvětlit, že tím, že tam dají to audio, tak zvýší pravděpodobnost toho, že se to audio bude zobrazovat v těch doporučeních. Pokud tohle budeme implementovat, tak o tom určitě musí vědět, aby dohlídli tam, že jejich kroky způsobí něco na mujRozhlase. Musí vždycky vědět, co způsobí to, že něco udělají.

Budete nějak informovat posluchače o tom, že ty rekomenační algoritmy používáte? Dozví se to nebo se to bude považovat za přirozený, že teď mujRozhlas funguje trochu jako Spotify?

Když jsi přihlášená na Drupalu, tak my u každého boxu pro webeditora máme takový otazník a ty když na něj klikneš, tak tady vidíš, co je to za box, podmínky a co do něj padá za obsah. Mně by se líbilo, kdybychom takovéhle prvky udělali i na ten mujRozhlas. Ve chvíli, kdy tenhle box je roboticky vybírán, tak tady by bylo nějaké íčko, na něj klikneš a zobrazí se ti „toto je obsah, který pro vás vybrala umělá inteligence nebo doporučovací mechanismus nebo robot a byl vybrán podle těchto pravidel“.

Takže ve chvíli, kdy budu mít účet i jako uživatel, tak se dozvim ty podrobnosti?

To je naše úvaha. Nám by se to líbilo. Mně dává smysl ta transparentnost. Stejně jako u Midjourney, když použiješ obrázek, který byl vyrobený AI, tak o tom musíš ty lidi informovat, na to máme i politiku a striktně to dodržujeme a líbilo by se mi, kdybychom to dělali i u toho doporučování. Například u těch boxů na homepage Radiožurnálu to nemá smysl, tam psát „tohle vybírá člověk“ nebo „sem padá nejnovější obsah, který má v sobě video a sem padá nejnovější obsah, který má tag sport“. To podle mě nikoho nezajímá, ale u toho doporučování je to podle mě zajímavá. Takže bychom tam dali nějaký disclaimer: „Tento obsah byl vybrán algoritmem, který zohledňuje relevanci pro vás podle vaší historie, relevanci pro vás podle vašich oblíbených audií, relevanci podle toho, co ostatní poslouchali k audiím, která posloucháte vy, a proto vám robot vybral tento obsah“. Je to úvaha, že bychom to chtěli udělat v rámci transparentnosti, ale ještě nevíme, kdy a jak přesně, ale víme, že chceme být transparentní.

Má podle tebe používání těchto prvků opodstatnění v médiu veřejné služby?

Určitě ano. Médium veřejné služby musí být ze zákona dobrý hospodář, takže musíme zohlednit i ty finanční aspekty. Používat stroj na výběr bude třeba levnější než to všechno nechat dělat editora. Určitě to má benefity i pro toho uživatele, když se vyvarujeme těch negativ jako uzavírání do bubliny. On si vlastně platí službu veřejnosti koncesionářskými poplatky a my třeba za jeho peníze vyrobíme něco, co by ho zajímalo a on se o tom nedozví, takže my část těch koncesionářských poplatků utratíme za doporučovací mechanismus, čímž docílíme toho, že pro toho uživatele a posluchače je větší šance, že se k tomu obsahu dostane. Já si myslím, že je to naopak naše povinnost a právo toho posluchače, že on nám platí peníze, abychom vyráběli obsah, který by ho měl zajímat a který ho zajímá a naše povinnost je, ho k němu dostat tak, aby si ho mohl komfortně poslechnout a nemusel se probírat, když to řeknu hnusně, balastem obsahu z 24 stanic a hledat, co by ho možná mohlo zajímat. My jsme povinni mu to nějak hezky podat, aby si to našel a pomohli jsme mu zorientovat se v tom obsahu, který pro něj produkujeme za jeho peníze. Takže si myslím, že je to dokonce naše povinnost.

Respondentka Anna Vošalíková, datum 16.10.2023

Jaké nástroje nebo prvky AI už využíváte v rámci aplikace mujRozhlas?

Co se týče technologických věcí, určitě ti Honza vyprávěl většinu z toho, ať už je to třeba rozpoznávání mluvicích nebo přepisovací systémy, tak tam je v nějaké kapacitě už nasazená umělá inteligence, byť samozřejmě ta definice je taková vágní. Obsahově to využíváme v několika projektech, ten asi nejobsáhlejší projekt se jmenuje Digitální spisovatel, který jsme začali dělat v roce 2020.

Je to projekt, který spočívá v tom, že umělá inteligence, konkrétně autoregresivní jazykový model GPT 2 až GPT 4, píšou povídky pro Český rozhlas. Začali jsme s tím v roce 2020, a to konkrétně právě proto, že jsme používali umělou inteligenci v nějaké kapacitě a chtěli jsme vědět, co by se stalo, kdybychom ji využili opačně, to znamená ne technologicky, ale úplně kreativně.

Další způsob, jak ji využíváme, je pomocí generování obrázků, to znamená nějaké generativní modely, které fungují při generování vizuálního obsahu. Tam pracujeme zatím ve velice malém měřítku. Používáme to třeba právě u digitálního spisovatele nebo u nějakých menších obsahových projektů. Tam řešíme ještě označování takového obsahu a tak dále. Potom samozřejmě obsahově jsme ji také použili v projektu Gott navždy, což je asi takový nejpracovanější projekt, který využívá umělou inteligenci, konkrétně syntézu hlasu, kdy jsme vlastně vysyntetizovali hlas Karla Gotta a čtení jeho autobiografie.

Jaké ty projekty měly smysl? Měly spíš popularizovat AI? Co je tam podstatné, proč jste s rozhodli je dělat?

U digitálního spisovatele tam byla chuť nějak experimentovat s umělou inteligencí, protože tehdy to fungovalo na základě takového jako jednoduchého API, byl to takový jako velmi rudimentární způsob komunikace s umělou inteligencí. My jsme se inspirovali u projektu, který se jmenuje *Digitální filosof* Dity Malečkové a Honzy Tyla a který se nám velmi líbil a chtěli jsme něco podobného využít pro Český rozhlas, protože jsme chápali, že je tady vzrůstající trend práce s umělou inteligencí a chtěli jsme nastartovat obecnou debatu a být u začátků debaty o kreativním využívání umělé inteligence, protože máme prostě pocit, že i veřejnoprávní médium má v tomhle nějakou roli a mělo by nastavovat ten diskurz nějakým způsobem. Takže u digitálního spisovatele to bylo tohle. Když jsme ten projekt potom rozvíjeli, tak jsme to chtěli po každé té sérii někam posunout, abychom nedělali jenom obsah pro obsah, protože lidských spisovatelů, kteří píšou skvělé věci je spousta. Takže první série byla čistě napsaná umělou inteligencí. U druhé série jsme vlastně navrhli kolaborativní modely, které replikovaly kreativní procesy skutečných lidských autorů, a tak jsme je do toho vlastně taky zapojili, protože oni spolupracovali s tou umělou inteligencí. Autoři jako Petr Stančík, Ondřej Neff, Bianca Bellová, řada dalších, Hana Lehečková, Pavel Bareš a František Kotleta spolupracovali s umělou inteligencí a výsledné povídky jsou synergie mezi nimi. U třetí série, která vznikla letos v roce 2023, jsme chtěli nejenom zase psát povídky pomocí chatu GPT, který vlastně až si troufnu tvrdit s loňským rokem a letoškem získal takovou tu masovou popularitu toho, že už to není API, ale je to chat nebo chat bot, se kterým se dá normálně mluvit. Tak jsme to použili nejenom k psaní tradičních povídek, ale i k nějaké reflexi rozhlasové práce. Nechali jsme si vygenerovat rozhlasovou reportáž, kterou načel Honza Pokorný a chtěli jsme vědět, jestli je třeba tohle nějaké využití, které bychom mohli do budoucna implementovat. Nechali jsme si vygenerovat dva díly podcastové série, kterou budeme potom celou spouštět začátkem příštího roku. Jmenuje se Rozluštěně záhady a je to

true crime konspirační teorie napsaná umělou inteligencí a chtěli jsme vlastně vědět, jestli můžeme být nahraditelní. Zjistili jsme, že úplně ne, protože pořád ten jazyk není až zas tak úplně optimální, ta témata samozřejmě nejsou optimální, pak jsou tam miliardy factcheckingových problémů, které prostě v umělé inteligence jsou. Celkově si myslím, že ta strategie, proč využíváme umělou inteligenci v obsahových projektech, je taková, že chceme jednak být u té diskuze, která prostě u veřejnosti probíhá a chceme s tím experimentovat. Český rozhlas jako médium, jakým je, má přicházet s obsahovými a technologickými inovacemi a chceme taky vědět, co to umí a chceme experimentovat, protože máme technologie rádi.

Jaké výzvy, třeba etické, vám to přinese?

Přineslo nám to strašně moc výzev a chtěli jsme... nebo já jsem laškovala s ideou, že bychom třeba u té první série posluchačům neřekli, že to napsala umělá inteligence a zeptali bychom se: „co si o tom myslíte?“ a pak bychom jim to řekli a viděli, jestli ta reakce je třeba jiná, jestli třeba soudí ten text negativněji na základě té umělé inteligence. Ale to jsme nakonec neudělali a rozhodli jsme se být v tomhle maximálně transparentní, protože jsme si vědomi toho, že naprosto zásadní devizou pro Český rozhlas je samozřejmě důvěryhodnost a nechtěli jsme s tím experimentovat tady v tom ohledu. To znamená, když máme nějaký obsah generovaný umělou inteligencí, ať už je to literárně-dramatický obsah jako v případě Digitálního spisovatele nebo je to syntetický hlas nebo je to nějaká obsahová záležitost, tak jsme si velmi vědomi toho, že musíme posluchačům dávat najevo, že to je generované, že vlastně do jisté míry nějakým způsobem interagují s umělou inteligencí, protože máme pocit, že to je naprosto zásadní. Takže to je třeba taková velká etická výzva, jak označovat obsah, který generuje umělá inteligence, vizuální obsah, audio obsah. Jestli je to nějaká znělka, jingle, jestli to stačí napsat do textu, jestli to musí zaznít v tom audio, že to je umělý obsah, nebo nějaké syntetické médium nebo jak s tím pracovat. Velká etická výzva třeba u digitálního spisovatele byla ta, a kterou nám potom posluchači taky omlátili o hlavu, že „bereme práci spisovatelům“. Což si myslím, že úplně nebereme, protože ty texty, když si to kdokoliv poslechne, tak zjistí, že nejsou srovnatelné s textem, který napíše člověk. Ani u té nejnovější série to není srovnatelné, takže tam spíš šlo o nějaké nastolování debaty, kterou se nám tedy velmi podařilo rozproudit, což mám radost, ale to bylo třeba taky etické dilema, jestli opravdu nebereme lidem práci, protože tady pětidílnou povídkovou sérii napsala umělá inteligence za dvě hodiny oproti tomu, že by se s tím nějaký lidský autor psal dlouho, dal by do toho prostě nějaké zásadní kreativní emoce... Stejně tak podobným dilematům čelíme i při generování toho vizuálního obsahu. Jestli nebereme práci někomu, kdo by to navrhl, nakreslil lépe, ale zatím ta strategie je taková, že to generování toho vizuálního obsahu používáme minimálně a většinou ještě ve spolupráci s nějakým grafikem, který to poladí, pošteluje, udělá, co musí, aby to nevypadalo nějak extra divně. Takže to jsou takové 3 asi základní největší výzvy, se kterými jsme se setkali.

A co chystáte teď do budoucna? Budete pokračovat v tom digitálním spisovateli?

Asi ano, pakliže se nám podaří pro tu čtvrtou sérii najít nějakou zase zásadní inovaci. My to právě chceme dělat tak, aby každá ta série přišla s něčím opravdu velmi jiným, protože si myslím, že nikdo neshledává žádný zásadní merit v tom, že jenom generujeme tuny textu a pak je dáváme do audia, protože by jako uměl každý. Takže pakliže se nám podaří přijít s

něčím, co je nějakým zásadním způsobem inovativní, nebo pokud ty jazykové modely třeba pokročí natolik, že bude zase záhodno s nimi experimentovat takovýmhle způsobem, tak určitě jo. Zmínila jsem ty Rozluštěné záhady, což je teda podcastová série, která vyjde, myslím si, že v lednu, únoru příštího roku. Je to desetidílná podcastová série, kde je každý ten díl věnován jednomu zločinu nebo konspirační teorii. A všechno je to kompletně syntetické. Nic z toho není pravda, takže my si chceme myslet, že to je takový první skutečně etický true crime podcast, protože ty crimy nejsou true.

A to řeknete těm posluchačům?

Ano, určitě. Ten zatím pracovní podtitul je *Zločiny, které spáchala umělá inteligence*, což je samozřejmě velmi provokativní, protože nikdo žádné zločiny nespáchal, ale čistě marketingově se nám to líbí, že by to mohlo přitáhnout. Zvažovali jsme dokonce, že bychom to nechali načíst i nějakým syntetickým hlasem, ale to už by bylo asi moc, takže budeme mít skutečnou lidskou moderátorku, která má nicméně naprosto vysyntetizovanou osobnost. Takže jméno je fiktivní, jaká je, jak mluví, to jsme si nechali randomizovat, ale bude to člověk, který tam pořád dodá tu emoci, takže to bude trošku uposlouchatelnější, si myslíme. To je takový velký obsahový projekt s umělou inteligencí, který právě teď chystáme a myslím si, že na to budou ohlasy. Kdybych mohla doufat, jaké na to budou ohlasy, tak to bude uposlouchatelné a bude to zábavné. Myslím si, že asi dostaneme trošku čochku za to, že neřešíme třeba skutečné zločiny, ale zase máme jiné podcasty, které se věnují true crime. Takže každý si přijde na své, ať už chce poslouchat opravdové zločiny naše nebo opravdové zločiny cizí anebo fejkový zločiny.

Ted', jestli se můžeme zaměřit na ty doporučovací algoritmy... co podle tebe využívání těch doporučovacích algoritmů přinese posluchačům, čtenářům za výhody a etické výzvy?

Výhody upřímně asi zásadní, když se podíváme i do zahraničí, jak fungují různé aplikace, ať už třeba BCC Sounds nebo Yle nebo jakékoliv jiné, tak víme, že personalizace je samozřejmě obrovské téma a čím personalizovanější, čím lépe umí ta aplikace doporučovat, co toho posluchače bude zajímat, tím je to prostě zkrátka lepší. Zásadní výhoda je, že se posluchač dostane k naprosto konkrétním tematickým okruhům, které ho zajímají. Nemusí to být vlastně ani tak personalizace tematická jako třeba nějaká geolokace, takže může dostávat extrémně aktuální dopravní zpravodajství nebo extrémně aktuální zpravodajství z regionu, které ho zajímá. Takže ty výhody si myslím, že jsou jednoznačné. Nevýhody pro běžného uživatele si na první pohled nedovedu moc představit, že by tam něco takového bylo. Může se možná, a to asi je docela často vytýkáno právě těm doporučovacím systémům, že člověka uzavírají do tematické bubliny nebo názorové bubliny. Že pakliže se mi líbí a jsem zastánce nějakého postoje, tak mi to spíš bude nabízet více toho postoje, než aby mi to nabízelo nějakou názorovou pluralitu, což může být třeba u média veřejné služby trošku problém, protože to může potom působit, že se k tomu posluchači dostávají informace jenom z toho jednoho spektra. Na druhou stranu, pakliže tohle bude dobře ošéfované, a myslím si, že dramatická většina obsahu Českého rozhlasu a pořadů Českého rozhlasu tu pluralitu mají v sobě zabudovanou, málokdy máme jenom pravičácký nebo jenom levičácký pořad, tak se tomu snad vyhneme. Ale obecně ty algoritmy, je to krásně vidět třeba na sociální síti X, že opravdu nabízí tak velmi intenzivně z jednoho spektra, že člověk má potom pocit, že mu všechno hraje do karet, i když velmi často se jenom nedostává k nějaké části těch názorů, které mu ten algoritmus nenabízí. Takže to je asi taková největší etická rezerva, které věřím,

že se třeba v rámci aplikace mujRozhlas podaří vyhnout, ale samozřejmě může se stát, že ne.

Co podle tebe využívání těchto prvků přinese pracovníkům rozhlasu? Zase nějaké výhody, nevýhody a výzvy?

Tam si myslím, že asi nejzásadnější výhoda je, že to nebudu muset editovat a doporučovat já, to mi přijde docela fajn.

Takže tvojí práci to nějak ovlivní přímo?

Ano, určitě. V tuhle chvíli je doporučování na mujRozhlas vlastně trojího způsobu. Jeden z nich je tedy už doporučovací systém, který nějakým způsobem funguje, ale pak tam máme takzvané klastrové doporučování, což je doporučování pořadů z jednoho tematického celku. Třeba když mám četbu na pokračování, tak pak to doporučuje Hru na sobotu, Hru na neděli, četbu na pokračování z jiné stanice, Počteníčko, Plk na nedělo a takové věci z toho literárně-dramatického žánru. Pakliže bude ten algoritmus dostatečně flexibilní v tomhle, tak tohle bude moct nastavit a nebudu to třeba muset hlídat já, bude se to moct hlídat podle tagů. Nebudu já muset hlídat, když přibude nový literárně-dramatický pořad, abych ho ručně zařadila do klastru, ale bude se to třeba moct vzít podle tagu „lit dram“ a rovnou to tam spadne a bude to doporučovat třeba z toho.

Druhá věc, určitě si myslím, že pracovníci rozhlasu, kteří to nepoužívají, ale vyrábí ten obsah a konzumují ho třeba právě v tom našem prostředí, tak asi budou předpokládám rádi, že kvalitní, pakliže bude kvalitní, doporučovací systém lépe odhaduje, co se k čemu hodí, co by posluchači, kteří chtějí jednu věc, mohli chtít i jako druhou. Dovedu si představit, že v tomhle ohledu to bude asi super, ale primárně si myslím, že pro mě nejzásadnější bude asi to, že to ovlivní moji práci, která bude do jisté míry automatizovatelnější a myslím si, že asi přesnější, protože já neuhlídám všechno.

Takže nemáš pocit, že tě to jako připraví o práci, ale naopak, že ti to třeba uvolní kapacity?

Určitě si myslím, že to tak bude. Myslím si, že by málokdo nějak zásadně lpěl na tom, že chce editovat takovéhle hordy materiálu. Hlavně si myslím, že to dává prostor pro více třeba editorské činnosti na jiných oblastech. Že už nemusí člověk dělat to doporučování a může dělat třeba nějaké zajímavé tematické shluky, které se potom dají nabízet třeba plošně, nebo se dají dělat tematické shluky, které se dají potom nabízet na základě jiného typu personalizace. Myslím si, že je to výrazně třeba kreativnější editorská práce než čistě mechanicky jenom hlídat, jestli mi do všech políček padá všechno, co mi tam má padat. Myslím si, že mi to tu práci usnadní a umožní mi to dělat více kreativní editorská rozhodnutí, která třeba nemusí úplně nutně reflektovat tu širší nabídky Českého rozhlasu, ale můžou reflektovat nějaký tematický zájem, což mi přijde, že je strašně fajn.

A jak teď vypadá tvoje práce v tomhle ohledu? Jestli tomu dobře rozumím, že ty fakt musíš projíždět a sledovat všechnen ten obsah?

U těch klastrů tam mám v tuhle chvíli nastavené čtyři defaultní klastry, kam dávám pořady. Mám tam *podcastový*, *sportovní*, *dětský* a *literárně dramatický*, a to musím nějakým způsobem updatovat. Pakliže je třeba nový podcast, to jsou třeba Buchty na Radiu Wave, a

oni mají velmi často nové podcastové produkty, tak ty musím házet do toho *podcastového*, aby se mi to doporučovalo tam. Tam je totiž několik úrovní doporučování a já tě nechci mystifikovat, ale myslím si, že u detailu audia, tam je doporučovací systém, ale u detailu pořadu tam je právě ten klastrový systém. Takže vlastně když jsi na detailu pořadu a pak tam máš výpis těch jednotlivých epizod, tak dole máš doporučování a já mám za to, že tam právě je to klastrové doporučování. To znamená, když jsi třeba na Hostu Lucie Výborné, tak tam dole ti to nabídne Blízká setkání, Buchty, Kriminálku a nějaké penzum, těch jakoby top pořadů. Když jsi na literárně-dramatické četbě na pokračování, tak ti to nabídne Četbu s hvězdičkou, Hru na sobotu, Hru na neděli, Počteníčko a ještě něco. A to je potřeba updatovat ručně. Takže v případě, že něco z toho už není aktivní, jako třeba Pokračování za pět minut, to byl pořad na Vltavě, tak je potřeba to ručně smazat a nahradit něčím jiným. Není to jako žádná zásadní fatální práce, jenom je potřeba mít na zřeteli, že tohle je nějaká administrace, kterou je potřeba hlídat, aby tam do toho nepadaly nesmysly, nebo aby to nedoporučovalo půl roku starý podcast, který už je nahrazený něčím jiným.

Budete editory nebo novináře nějak připravovat na práci s těmi algoritmy?

No v tuhle chvíli má mujRozhlas jednoho editora, a to jsem já, takže já se na to připravím, nicméně samozřejmě s příchodem mujRozhlas 2.0 budou mít do editace přístup webeditoři všech stanic, kteří tam budou editovat svůj obsah a v případě, že tam bude nějaká nová implementace nějakého doporučovacího systému, tak samozřejmě bude potřeba jim to vysvětlit a připravit je na tu práci s tím. Dovedu si představit, že s nějakým „paušálním systémem“, který by byl do jisté míry automatizovaný, tak to penzum lidí, kteří s ním budou pracovat přímo, bude asi menší. Nebude to každý web editor, ale bude to patrně asi hlavní webeditor a možná ještě jeden nějaký jakoby editor obecný. A myslím si, že u těch webeditorů stanic bude primárně, aspoň v těch prvních fázích, velká zodpovědnost hlídat, jestli to doporučování je aktuální, jestli to doporučuje tak, jak to má. Protože oni jsou samozřejmě nejlépe obeznámeni s obsahy těch stanic. Tam si myslím, že bude i potřeba reportovat tu jejich personalizaci, jestli třeba to opravdu je něco, co oni sledují. Takže v té první fázi si myslím, že to bude spíš takové jako učení na všech stranách. Nicméně samozřejmě, už jsem mluvila o té transparentnosti, i tady v tomto ohledu bude potřeba, aby editoři chápali, jak ten systém alespoň elementárně funguje, aby byli schopni říct, jestli jim to doporučuje kvalitně nebo nějak divoce nebo jestli je to úplně mimo, což se nám taky stávalo na začátku s tím, co máme teď. Takže určitě s tím budeme pracovat.

Jo, takže spíš s webeditory jednotlivých stanic? Protože já potom v té diplomce mám rozhovory s editory z iRozhlasu, tak těch se to třeba bude týkat? Nebo třeba zpravodajství?

Myslím si, že editoři zpravodajství jako takoví budou jako ještě o jeden stupeň více detašovaní od té práce s tím algoritmem. Z větší části proto, že ti většinou mají dost práce s tím, aby ohlíželi, co v tom zpravodajství zazní. Primárně v tom zpravodajství. Jestli je to vypublikováno, jestli text, který je k tomu přiřazen, je adekvátní, jestli titulek není

clickbaitový, jestli prostě tohle je vlastně v pořádku, takže si myslím, že u nich je ještě o něco méně času na to zjišťovat, jestli ten doporučovací algoritmus funguje tak, jak má. Myslím si, že jejich velkým cílem nebo jejich velkým úkolem bude asi možná právě sledování těch bublin. Jestli nám to nesklouzává nějakým způsobem, jestli to, že se u nás konzumuje více obsah, který je zaměřený na názorovou pravici, nevede k tomu, že se ten systém celkově učí, že má doporučovat víc tohle a méně třeba nějaké levicové názory nebo levicové obsahy nebo obsahy s nějakým levicovým přesahem. To je možná na nich velká zodpovědnost a může se stát, že se ten systém naučí, že tím, že se u nás konzumuje více tenhle obsah, tak ho doporučuje víc. Takže si dovedu představit, že třeba tohle bude velká role pro ně, aby to sledovali. Ale myslím si, že ten jejich day to day opravdu nebude sedět hodiny a hodiny před doručovacím systémem a nějak zjišťovat, jestli doporučuje správně, ale možná si to myslím úplně špatně...

Takže asi předpokládám, že budou probíhat nějaká třeba školení uzpůsobená jejich potřebám?

Určitě bude. Dovedu si představit, protože jsou samozřejmě zodpovědní za ten obsah ve velké míře a je to jejich hlava na špalku, že budou chtít vědět, jak to funguje. Na základě jakých premis se to doporučuje a všichni jsme si v zásadě vědomi nějakých machine learning bias a těhle věcí, takže si myslím, že to bude pro ně taky relevantní vědět, jak tomuhle čelíme, jestli tomu nějak čelíme a nebo jestli to budeme jenom sledovat a ručně třeba nějak upravovat podle toho, jak to budeme vnímat, jak to budou vnímat oni, protože oni jsou za tu názorovou pluralitu zodpovědní, výrazně víc než webeditoři.

Budete informovat čtenáře o tom, jakým způsobem a jestli vůbec používáte ty doporučovací systémy?

Pokud zase vezmu příklady ze zahraničí, tam moc není upozorňováno na to, že ten obsah je doporučován automaticky. Většinou je to tam jenom jako *doporučení pro vás* a tam už bude to konkrétní personalizované doporučení. Já si myslím, že to může být součástí nějaké kampaně, která třeba doporučuje lidem, aby se přihlašovali v prostředí mujRozhlas, protože pouze přihlášení budou mít tyto výsostně personalizované věci. Nepřihlášení uživatelé budou mít samozřejmě nějaké spíš editorské doporučení nebo doporučení „co právě frčí“ z analytik. Takže dovedu si představit, že to může být součástí nějaké komunikační kampaně typu „Jestli chcete opravdu personalizovaný obsah přímo pro vás na základě algoritmického doporučení, tak se musíte přihlásit, protože pak vás můžeme monitorovat, sledovat, můžeme znát vaše zvyklosti líp a můžeme podle nich potom doporučovat“.

Takže v téhle kapacitě určitě ano, protože lidi asi budou chtít vědět, proč tady o paní Voňáskové víme, že chce sledovat S Jarmilkou na kafi a my jí řekneme „Jo, tak protože my víme, že to máte ráda, tak vám to tady algoritmus doporučil.“ Ale nedovedu si představit, že by třeba v tom prostředí bylo nějaké políčko *toto doporučuje algoritmus* nebo *toto je generované* nebo něco takového. Nebývá to zvykem a myslím si, že v tom není žádné mystérium. To není žádná přidaná hodnota, když člověk ví, že mu to nedoporučuje člověk, ale doporučuje mu to počítač. Na rozdíl třeba od těch obsahů, kdy si myslím, že je výsostně důležité, aby člověk věděl, že konzumuje syntetické médium v jakékoliv kapacitě.

Má podle tebe používání těchto nástrojů nebo prvků opodstatnění v médiu, veřejné služby, proč ano, proč ne?

Stoprocentně má, už jsme to vlastně zmínili. Vyplývá to ze všeho, co jsme řekli. Jednak to, že umožňují usnadňovat práci. Člověk se může věnovat více kreativním pracem nebo i

kreativnímu editorství, což mi přijde, že je vlastně často opomíjený segment. Dále je to zvykem v médiích veřejné služby mít kvalitní doporučovací systémy. Je zvykem v médiích veřejné služby, aby technologicky držela krok s dobou nebo byla alespoň nějakým způsobem inovativní. Takže určitě ano. Nemá to smysl, pakliže k tomu to médium veřejné služby přistupuje neopatrně, nebo pakliže nevěnuje dostatek práce právě těm etickým otázkám nebo tomu vývoji a rozvoji.

Co se týče konkrétně Českého rozhlasu, tak si myslím, že to má určitě svoje místo v nějakém jako UX designu. Třeba u mujRozhlas 2.0 si nedovedu představit, nebo respektive vím for fact, že v Českém rozhlase se tyto věci neberou na lehkou váhu, takže si myslím, že budeme schopni dodat posluchačům něco, co je velice solidní.

Respondentka Dalia Moudrá, datum 17.10.2023

Jaké nástroje nebo prvky AI Český rozhlas používá?

Nejdřív to vztáhnou na iRozhlas jako takový. Na iRozhlasu si tak malinko pohráváme s doporučováním, to znamená s nějakým doporučovacím algoritmem. Potom, co využíváme nějak jako institucionalizovaně, jsou přepisy, to znamená přepis audia do psaného textu a teď začínáme pracovat i s tou druhou cestou, to znamená přepis psaného textu do audia. Ale to je pro vnější uživatele i pro redakční uživatele zatím dost mlhavá představa. Protože si s tím hrajeme spíš my produktůvci s vývojáři a máme v plánu to uvést do provozu v blízké době. Ale co se týče redakce, tak tam teprve vůbec pracujeme na tom procesu, jak by ta syntéza měla fungovat i ve vztahu právě k redakci. To je syntéza. Potom se zabýváme třemi věcmi. Děláme experimenty s různými volně dostupnými nástroji. Pro iRozhlas je to hlavně ChatGPT nebo další jiné textové modely, kdy si hrajeme s tím, jak nám to může pomoci dejme tomu při generování titulků nebo při generování postů na sociální sítě a hrajeme si i s tím, aby ten ChatGPT... kdy já teď říkám ChatGPT, ale právě Honza Cibulka si s tím pořád různě hraje a třeba mění různé modely, které to vlastně používá, ale když o tom mluvíme v redakci, tak říkáme ChatGPT.

Takže to není ten ChatGPT?

Tak oni ty generativní modely jsou různé. Určitě vím, že to začínalo na ChatuGPT, ale mám pocit, že si tam třeba hraje i s nějakými jinými modely, ale to ti klidně potom konkrétně řekne on. Nicméně konkrétně tento nástroj používáme experimentálně. Není to nijak zavedené vyloženě do infrastruktury. Je to o tom, že kdo z těch redaktorů třeba k tomu má blíž, tak to může využívat ke své práci. Kdo k tomu nemá úplně tak blízko, tak třeba ví, že to existuje, ale nepoužívá.

To je tedy ChatGPT. Co se týče iRozhlasu, tak zatím nepracujeme s Midjourney, ale obecně se tady v rozhlase o jeho využití bavíme... nebo zase obecně nějakého modelu, který umí generovat obrázky a máme tady pro to nějaké způsoby využití, které bychom rádi vyzkoušeli. Zatím se to rozjíždí, ale u zpravodajství je docela problém u fotek, protože v okamžiku, kdy je to zpravodajská fotka, tak by to ideálně měla být opravdu reálná fotka. V okamžiku, kdy je to ilustrační snímek, tak jsme zatím asi úplně nenašli tu velkou potřebu, kdy by nám to natolik usnadnilo práci, protože už tady přece jenom máme nějaký archiv ilustračních snímků a záběrů. Máme vlastní fotografy, kteří jsou schopní nám cokoli vyfotit, když je o to poprosíme, a tím pádem se v rámci toho zpravodajství, které musí být pro uživatele a pro čtenáře vnímáno jako spolehlivé, úplně nehrneme do toho zkoušet obrázky generované umělou inteligencí. To je tedy asi tak těch pět hlavních oblastí, které řešíme.

Ještě mě rovnou napadá k tomu ChatuGPT a podobně, jak jsi říkala, že ho někdo používá a někdo ne... tak máte na to třeba nějaký kodex nebo pravidla?

To je přesně věc, kterou hodně řešíme. Na co máme interní směrnici je syntéza hlasu. To je vyloženě rozpracované. Jak se s tím může pracovat, jak se s tím nemůže pracovat a jsou tam vypořádané i vztahy s poskytovatelem toho hlasu, což je reálná osoba žijící nebo i nežijící. V okamžiku, kdy jsme řešili syntézu Karla Gotta, tak to je nežijící osoba. V té směrnici se s tím pracuje. Jakým způsobem se s takovým hlasovým modelem nakládá. Syntéza, kterou teď řešíme na iRozhlasu a chceme ji v blízké době spouštět, je hlasový model založený na hlasu žijící osoby. Tam se s tím zase pracuje malinko jinak, ale je to v té směrnici jasně

popsané. To je normálně směrnice, která je dostupná na intranetu. Další směrnice, která existuje, je na Midjourney... já říkám Midjourney, člověka to svádí k tomu říkat ty názvy značek, ale samozřejmě ty datové modely nebo ten princip je prostě generativní umělá inteligence, která pracuje na bázi obrázků. Když mluvím o Midjourney, tak myslím toto. Tak na to už máme taky směrnici, která vychází tady z potřeby hlavně Aničky Vošálíkové. Protože ta má nějaké use cases, kde by s tím chtěla víc experimentovat. Na rozdíl od nás v tom zpravodajství, kde je to, jak jsem říkala, trochu problematičtější. Ale ta směrnice obsahuje i zpravodajství a to, jak by se s těmi výstupy mělo pracovat. Ve všech případech, kdy používáme umělou inteligenci, tak vždycky interní směrnice říká, že uživatel musí být srozuměn s tím, že je to výstup umělé inteligence. To znamená, musí to být dostatečně vizuálně nebo audiálně jasně označeno u toho díla. Pokud to je například ta hlasová syntéza, tak tam pracujeme s dvěma typy. Když to je literárně-dramatické dílo, tak to opravdu musí být přímo úplně u toho audia oznámeno, že: „Teď budete poslouchat něco, co vzniklo umělou inteligencí“. Pokud se u syntézy hlasu bavíme o zpravodajském obsahu, tak tam v té směrnici, pokud se nepletu, je stanoveno, že to musí být dostatečně vizuálně označeno na místě, kde ten uživatel kliká na tlačítko *play*, když si začíná přehrávat to audio. To samé u obrázků, tam jsme poměrně dlouho řešili, jakým způsobem to označovat, aby to bylo jasné. S tím, že jsme i řešili, kdo je autorem toho obrázku. Protože je rozdíl být „autor obrázků“, což je vlastně Midjourney. Ale tady se spíš řeší autorství toho promptu a autor toho promptu je zase ošetřený v té interní směrnici: kdo to může být, kdo to nemůže být, jakým způsobem se to zdrojuje.

Co se týče ChatuGPT, tak tam momentálně ještě směrnice neexistuje, ale je v přípravě a nebyla ještě vydaná. Bude to ale mít stejné principy typu toho, že pokud by se uživatel dostal na článek, který je napsaný ChatemGPT, tak musí jasně být v úvodu tohoto článku uvedeno, že tento článek nepsal člověk, ale psal ho nějaký „ChatGPT“. Už je to ale trochu problematičtější, protože jako výsledek té práce, na rozdíl od Midjourney nebo syntézy hlasu, se používá přímo ten výstup, který vypadne z umělé inteligence. U ChatuGPT se velmi těžko identifikuje, co je ten výstup. Protože i pokud to nějaký redaktor použije, aby třeba sestavil tweet nebo titulek, tak je velmi pravděpodobné, že to používá spíše jenom jako nástroj pro brainstorming, a potom ten výsledný text, který se opravdu někde publikuje, je stejně malinko pozměněný. Co je důležité u všech typů výstupů umělé inteligence, ať už je to ta syntéza nebo obrázek nebo text, tak vždycky má rozhlas takové stanovisko, že to musí procházet kontrolou nějakého reálného člověka. Nikdy se nestane, že by se publikoval výstup té umělé inteligence, aniž by tam nebyl ten člověk, který se na to podívá a řekne: „Ano, takto to můžeme publikovat.“ Nebo aby to případně ještě poupravil, a pak to publikoval. Takže tohle je důležité. Obecně tady v rozhlasu máme takový projektový tým, který se právě zabývá implementací nástrojů umělé inteligence do struktury Českého rozhlasu nejenom právě do zpravodajství, nebo tady na digitálních službách, ale obecně i napříč IT a dalšími odděleními. V rámci toho týmu teď směřujeme k tomu, abychom vydali ucelený dokument, který bude popisovat, jak se Český rozhlas staví k práci s umělou inteligencí a k těm výstupům. Tenhle dokument by měl vyjít asi někdy do konce tady toho roku a myslím si, že potom bude veřejně dostupný, ale víceméně v tom dokumentu bude plus mínus to, co tady říkám, nebo co ti asi říkal i Honza Misák nebo Anička, protože všichni tři jsme součástí toho týmu. Takže tyhle myšlenky potom budou promítnuté v tom uceleném dokumentu.

Co ještě teď plánujete využívat?

My v tuhle chvíli asi nemáme nějakou věc, o které bychom věděli, že jí zatím nepoužíváme a nemáme ji vůbec zmapovanou a chtěli bychom ji začít používat, ale spíš se chceme zaměřit na zdokonalování toho, co už máme. To znamená to doporučování, syntéza, ty přepisy, anebo právě ta práce s tím ChatGPT. Už jsou to věci, které jsme si nějak jako osahali a chceme to zdokonalovat. Mám se teď zaměřit na to doporučování teď?

Jo, prosím.

Na iRozhlasu, jsme měli asi do června boxík *Robot doporučuje*, což byl vlastně boxík, který byl napojený na Recombee a na základě informací o uživateli, to znamená cookies a informacích o jeho prohlížení webu, jsme mu v tom jednom boxíku doporučovali obsah, o kterém Recombee vyhodnotilo, že by ho mohl zajímat. Když jsem iRozhlas přebírala někdy v únoru, tak jsem trošku zkoumala, jak tady tohle jako funguje a nebyly tam úplně nějaké zásadní výsledky, jako že bychom si díky tomu boxíku ověřili, že ano, toto je něco, co nám jako funguje násobně líp než ostatní patra na té homepagi, ale ta chyba byla na naší straně, protože ta implementace toho Recombee byla jenom úplně základní, a proto, aby to Recombee mohlo fungovat opravdu dobře, tak by bylo potřeba, abychom jim posílali mnohem víc dat než to, co jsme jim posílali. Na tom jsme začali pracovat, ale pak jsme bohužel zjistili, že jsou tam nějaké trošku větší technické problémy, takže jsme dočasně ten boxík úplně skryli někdy v tom červnu s cílem vyřešit ty technické problémy a nasadit ho v lepší podobě. Pak bohužel následoval hackerský útok, kdy jsme celé léto neřešili vlastně nic jiného, než abychom obnovili a navrátili i ta naše média, o která jsme přišli. Tím pádem se všechno ostatní posunulo trošku jako dál. Teď, řekla bych tak od půlky září, se k tomu vracíme. A to znamená, že teď už máme opravené ty technické problémy, ještě čekáme na nějaké malinké dodělávky, které jsou vlastně jako na třetí straně, ale já předpokládám, že třeba přes příští týden bychom ten boxík už měli navrátit. Bohužel opravdu tam byla ta mezera od června do asi konce října, kdy tam žádné to doporučování nebylo. Což mě osobně fakt mrzí... já jsem dlouhou dobu předtím pracovala na Seznamu, který má svoje úplně vlastní doporučování, které si vyvíjí uvnitř toho baráku, a to vím, že funguje velmi dobře a že na tom je ten Seznam jakoby postavený. Pak jsem pracovala chvíli na Primě a když jsem tam byla, tak se tam právě zavádělo Recombee, což znamená ne custom interní systém, ale systém třetí strany. Ale zase vím, že to Recombee může fungovat strašně dobře a vím, že to, co mi tady na iRozhlasu děláme, je zatím úplně jenom škrábání po povrchu toho, co bychom mohli z toho Recombee vytěžit. Takže já osobně se na to těším, až to tam navrátíme, ale zároveň vím, že to je jenom ta špička ledovce. Co budeme řešit příští rok ve vztahu k doporučování, tak je právě přesně to, abychom se dostali do fáze, kdy jsme si opravdu, že máme podchycené ty klíčové body na tom webu, a nejenom na webu, ale i v mobilní nebo televizní aplikaci. A kdy budeme ve stádiu, kdy víme, že „ano, uživatel u nás opravdu dostává obsah, který on chce vidět a o kterém víme, že ho bude zajímat“. U toho zpravodajství řešíme samozřejmě takové problémy jako jak neuzavřít toho uživatele do nějaké jeho bubliny. Protože samozřejmě jako zpravodajský web potřebujeme těm uživatelům servírovat aktuální zpravodajství. Ale myslím si, že ve výsledku to uzavírání do bubliny je taková obava, která je asi relevantní, ale zároveň osobně se toho až tolik nebojím, protože abychom někoho uzavírali do bubliny, tak bychom museli mít mnohonásobně větší množství obsahu, než na iRozhlasu máme a zároveň by musel být i natolik tematicky odlišný. Například to uzavírání do bubliny se dělo a je popsáno třeba na YouTube, kde jsou opravdu miliony videí na různá témata a uživatelům se stávalo, že jim YouTube doporučoval pořád jenom videa z toho tématu. Na iRozhlasu se osobně tady toho úplně neobávám, protože jak jsem říkala, prostě ta rozmanitost toho obsahu tam zase není až tak

velká, aby se mohlo stát, že najednou uživatel u nás bude číst pouze nějaké jedno téma a k ostatním se vůbec nedostane. Nicméně je to určitě aspekt, který řešíme, ať už tedy interně nebo s firmami, se kterými na tom doporučování spolupracujeme, tak vždycky tohle říkáme. U nás to není o tom, aby tam uživatel nutně strávil co nejvíc času čtením toho obsahu, ale... nebo samozřejmě ano, ale zároveň je to i o tom, aby opravdu ten uživatel k nám přišel a dostal ty aktuální zprávy, které jsou teď zrovna pro tuto chvíli relevantní. Takže tohle je aspekt toho doporučování, který si myslím, že je třeba jiný než na mujRozhlasu, protože na mujRozhlasu nám jde opravdu o to, aby tam ten uživatel strávil co nejvíc času a poslouchal co nejvíc audií. Já jsem navázala na to, co máme a už jsem rovnou přešla k tomu, co chceme mít. Důležitý aspekt je ten, že coby součást rozhlasu tak nemyslím si, že se nám stane v blízké nebo dohledné době, že bychom jeli čistě na doporučování, jako to třeba funguje na Seznamu. Seznam má na své homepage opravdu jenom omezené množství míst, které jsou redakčně spravované a ten zbytek, a je to větší poměr, je čistě doporučovaný. Nemyslím si, že by to teď na iRozhlasu byla úplně otázka nebo směr, kterým bychom chtěli jít, protože vždycky budeme chtít mít nějaký editorský dohled nad tou homepage. Ono to může znít trošku zvláštně, protože většina médií nebo většina obsahových webů řeší hlavně... nebo jejich hlavní landing page je detail článku, na který se uživatel dostane buď ze sociálních sítí anebo třeba v České republice často právě z homepage Seznamu. Pro nás je samozřejmě detail článku taky velmi důležitý, ale my musíme pracovat i s tím, že ta homepage je pro nás velmi důležitá, protože máme velké množství uživatelů, kteří k nám přichází opravdu napřímo tak, že si zadají iRozhlas.cz. Tím pádem, pokud je ta homepage to první místo, na které uživatel přijde, tak my vlastně vždycky chceme mít trošku dohled nad tím, co tam uvidí. Když budu jako úplně předpovídat, tak si myslím, že se to může proměnit, ale třeba v horizontu pěti let. Ale v tuhle chvíli, a ještě nějakou dobu to tak bude, je to stanovisko takové, že určitě homepage je vždycky editorsky spravovaná. Můžeme tam mít plochy, které jsou doporučované, ale ty plochy nebudou převládat nad editorsky spravovanými plochami. Takže to je něco, na co myslíme a čím se, si myslím, docela odlišujeme i od těch ostatních médií, právě proto, že nemáme až tak velkou motivaci vydělávat na reklamě, to znamená opravdu dávat tomu člověku do popředí titulky, které budou třeba až trošku clickbitové, aby na ně kliknu. Ale naopak, my jsme trošku ta protiváha tady tomu. Můžeme si dovolit servírovat těm uživatelům to zpravodajství v takové podobě, aby nebylo clickbaitové, aby bylo vyvážené a aby se ten člověk i jenom tím, že si projede tu homepage, dozvěděl opravdu, co se v tuhle chvíli děje v České republice a ve světě a co je důležité. Takže to jsem ještě chtěla zmínit ve vztahu k tomu doporučování... což ale neznamená, že bychom si neprohrávali s myšlenkami... a teď bych teda chtěla dát velký disclaimer, protože to je něco, co budeme chtít v příštím roce zkoušet, ale nevíme, jak to dopadne a bude se to určitě ještě proměňovat. Bude to série různých experimentů, ale máme v hlavě takový nápad, že bychom vedle naší homepage, která je editorsky spravovaná, postavili ještě sekci nebo stránku, která bude typu *pro vás* a ta bude opravdu doporučovaná. A pokud víme o tom uživateli, že má oblíbeného autora nebo oblíbené téma, tak právě tady na té stránce *pro vás* mu tady tyhle věci upřednostníme. Nikdy to nebude clickbaitové, protože nemáme clickbaitový obsah, ale už to bude malinko přizpůsobené tomu uživateli. A to si myslím, že je z mého pohledu největší výzva pro příští rok. Abychom tady tuhle stránku nastartovali a abychom pomocí různých experimentů došli k tomu, jak by mohla vypadat.

. A víme, že ať už se bavíme o sociálních sítích, nebo ať se bavíme třeba právě zase o homepagi Seznamu, tak prostě to, že já někde jsem a tam se mi opravdu nabízí obsah, který mě zajímá a můžu nekonečně scrollovat po té stránce, sem tam si něco rozkliknout a budu

pořád engaged, tak to si myslím, že je pro iRozhlas velká výzva a něco, co určitě chceme. Protože je pro nás důležité, aby ti uživatelé u nás trávili co nejvíc času, ale nejdeme na to úplně tak přes mrtvoly. Takže ta stránka *pro vás*, to je pracovní název, je něco na čem si myslím budeme v příštím roce úplně nejvíc pracovat. Druhá věc, na čem budeme pracovat ve vztahu k doporučení, je určitě doporučení na detailu článku, protože tam momentálně nic moc nedoporučujeme. Je tam nějaký úplně jako velmi hloupý algoritmus, ale ten vůbec nepracuje s preferencemi toho uživatele a s tím, co o tom uživateli už víme. Je to prostě tak, že když víme, že na iRozhlasu se aktuálně čtou tyto 3 články, tak je dáváme na ten detail toho článku *Mohlo by vás zajímat*, ale v tom okamžiku, my už neřešíme, jestli ten uživatel třeba ten článek už nečetl. Takže tohle je další část, co budeme v příštím roce řešit a optimalizovat a vůbec nastartovat to doporučení na detailu článku a pak ho samozřejmě optimalizovat. Třetí část, která je pro nás stejně podstatná, ale pro uživatele ne tolik viditelná, je doporučení v adminu nebo doporučení ne pro toho koncového čtenáře, ale pro redaktora, který píše ten článek. V tuhle chvíli fungujeme tak, že redakce iRozhlasu je poměrně malá, nebo těch lidí tam není úplně mnoho a jsou natolik zapálení, že opravdu mají docela přehled o tom, co se u nás vydává a mají přehled o tom, že pokud píšou článek o nějakém tématu, tak ví, že třeba 2 měsíce předtím o tom stejném tématu psal jejich kolega a jsou schopní si ten článek dohledat a propojit ho. Ale házeme jim tím jakoby klacky pod nohy, protože pokud si ten redaktor nevzpomene, že nějaký takový článek existuje, nebo se mu ten článek nedaří najít, tak je mnohem pravděpodobnější, že se na to prostě vykašle a nebude ho tam prolínkovat. Samozřejmě máme možnost vyhledávání podle nějakých štítků a rubrik. Ale s těma štítkama je třeba problém v tom, že někdo založí štítek *covid* a někdo jiný založí štítek *koronavirus* a v tu chvíli jsou články, které jsou potom přiřazené do štítku *covid* a nějaké jiné články do štítku *koronavirus* a jsou to vlastně články, které by měly být pod jedním štítkem. Ale dokud si toho nevšimne někdo... oni si toho většinou všimnou po nějaké době, že mají dva štítky na stejnou věc, ale pak je potřeba ty štítky sloučit, je potřeba všechny v redakci edukovat o tom, že oficiálně používáme třeba *covid* a je potřeba doufat, že nikdo jiný nepřijde za týden a nezaloží znova štítek *koronavirus*. Tohle se v tom množství lidí, kteří v té redakci jsou, dá uhlídat do jisté míry, protože jich zase není tolik. Ale na druhou stranu se nedá očekávat, že to bude vždycky úplně neprůstřelné, protože prostě je to jejich práce. Oni se snaží co nejrychleji vydat článek na téma, které je aktuální. Pokud se jim nedaří najít štítek *koronavirus*, protože se den předtím sloučil do štítku *covid*, tak ho založí, protože je třeba nenapadne, že teď máme štítek *covid*. A tohleto jsou věci z mého pohledu jako produkt'áka, které bychom jim měli usnadňovat a tyhle starosti by vůbec neměli mít ti redaktori, kteří jsou tady od toho, aby měli klid na svojí práci, na to psaní těch článků. Takže pracujeme s různými variantami, jak jim tohle usnadnit. Jedna celá část jsou ty štítky, ale to vlastně není až tak spojené s umělou inteligencí, ale druhá část, a co je umělá inteligence, je právě to, že my teď pracujeme na novém adminu pro iRozhlas. V rámci nového adminu chceme mít vyložene při editaci článku sloupec na pravé straně, který bude fungovat trochu jako ChatGPT a bude jim tam psát: „Vidím podle toho, co píšeš do těla článku, že je to podobné tady tomu článku, který už máme. Nechceš ho tam prolínkovat?“ Ten redaktor v tu chvíli získává už mnohem větší přehled o tom, jaké články máme v databázi, co případně může prolínkovat. Nebo i ve smyslu, aby třeba nepsal to samé, co už se jednou napsalo... což není reálné nebo nestává se to velmi často, ale i tady k tomu to může sloužit. „Aha, tak on už o tom kolega psal třeba před půl rokem, tak se podívám, co o tom psal.“ Tohle zapojení umělé inteligence je pro nás strašně důležité, byť ve výsledku ten koncový čtenář to vůbec nevnímá, protože ten prostě dostane článek s nějakými více či méně

relevantními souvisejícími články. Když říkám umělá inteligence, tak já to vnímám tak, že to funguje trošku jako ChatGPT, ale tam je to část ChatGPT, ale část je to doporučování, protože to je práce s tou databází a hledáním nějakých podobností mezi těmi existujícími články a mezi tím vznikajícím článkem.

A ten admin je teda řekněme program, ve kterém teď pracujete a vy ho jako předěláte?

My pracujeme v Drupalu, což je jeden z takových nejznámějších systémů na správu obsahu. Drupal nějak funguje, ale víme, že máme různé technické, ale právě i funkční požadavky, které už víme, že nejsme schopní do té současné verze toho Drupalu implementovat. Takže přecházíme na vyšší verzi Drupalu. Je to pořád stejný systém, ale vyšší verze, která nám toho umožňuje víc i z hlediska bezpečnosti nebo technologického rozvoje. A tím, že přecházíme, tak si ten admin budujeme víceméně od začátku, protože si říkáme, co už víme, že tam potřebujeme, co nám funguje v tom starém, naopak co tam nefunguje a snažíme se to udělat lepší. Já si myslím, že asi vy možná na Wavu pracujete taky s Drupalem ale s trošku...

Já se s tím nedostanu do kontaktu, ale už to chápu. Děkuju. Jak promění využívání těchto prvků práci editorů nebo novinářů a jestli si myslíš, že jsou na to připravení?

Připravení... no já musím říct o sobě, že já vlastně nejsem úplně takový ten šílený nadšenec... tak dejme tomu, že třeba rok zpátky, kdy začal takový opravdu velký boom těch ChatGPT a podobných nástrojů, tak v tu chvíli jsme samozřejmě všichni byli jako strašně nadšený a říkala jsem si: „Hele super a za rok už tady vůbec nebudou pracovat lidi, už to všechno bude dělat umělá inteligence.“ Ale máme o rok později, kdy jsme si mysleli, že už se to všechno bude dělat úplně jinak a osobně za ten rok, byť jsem do toho šla s tím nadšením a stále jsem nadšená z toho, co to umí, tak vlastně už tam taky nastal takový ten reality check v tom smyslu, že když potom člověk opravdu zkouší aplikovat ty dostupné nástroje, tak se dostává do situací, kdy zjišťuje, že ono to nějak jako funguje, ale ta práce toho člověka tam vždycky musí být. To znamená, když to dám na příkladu... Mě strašně nadchla varianta, že bychom měli třeba krátké shrnutí témat. Dejme tomu, že máme téma kutnohorská sekta, což je teď náš docela velký investigativní případ na iRozhlasu ve spolupráci s Radiožurnálem. Já jsem si říkala, že by bylo hezký, když o tom vyšlo třeba 20 článků, tak pro uživatele, který s tím přichází do styku poprvé, by bylo dobré aby dostal odstavec textu, který by řekl: „Kutnohorská sekta je tady o tom, tohle jsme zjistili, teď probíhá soud.“ Fakt krátké shrnutí. Velmi jsem teda uháněla právě i Honzu Cibulku, abychom to nějak dali dohromady, že to přece dává strašnej smysl. Tak jsme to zkusili spolu několikrát, různými způsoby a ve výsledku jsme došli k tomu, že ano, ono to umí napsat nějaké shrnutí... to shrnutí se nám i podařilo po různých pokusech a omylech, kdy nám to tam dávalo úplně nesmyslné věci, že se to naoko tváří, že je to shrnutí té kutnohorské sekty, ale pak tam ve třetí větě je zmínka o osobě, která tam nikde vůbec nefiguruje a vyloženě si to ten Chat GPT vymyslel. Tak sice se nám to z téhle fáze podařilo dostat do fáze, kdy to bylo fakticky správně, ale nebylo to zajímavé. Prostě ten ChatGPT to nikdy nepopíše natolik čtenářsky zajímavě, nevystihne ty nejdůležitější věci, jako to udělá ten redaktor, který je s tím případem srozuměný. Takže tady tímhle dlouhým úvodem jsem vlastně chtěla říct, že já se úplně nedomnívám, že bychom teď byli ve fázi, kdy se musíme reálně bavit o tom, že roboti nahradí lidskou práci. Myslím si, že ty nástroje můžou fungovat jako pomocníci pro lidi, ale úplně se neobávám toho, že by v dohledné době byli novináři bez práce. Takže ve vztahu k osobě jako redaktora nebo editora to vidím takhle. Co si myslím je ale to, že konkrétně třeba ti naši redaktori ty nástroje používají dost nahodile a někdo třeba k tomu má blíž, tak si už našel nějaký způsob, jak s

tím pracovat někdo, kdo k tomu nemá blízko, tak může klidně žít v tuhle chvíli, jako kdyby ty nástroje neexistovaly, nijak se to neodráží na jejich práci ani jednoho ani druhého, protože ta práce je vždycky o tom vypublikovat na iRozhlasu dobrý zpravodajský text. Ale někomu to třeba může usnadňovat práci i ve smyslu toho, že si udělá třeba brainstorming s ChatemGPT a potom na základě toho brainstormingu píše článek a píše ho už opravdu sám. Je to jeho autorský článek, ale předtím je tam příprava, kterou si udělá společně tady s tímto. Takže to si myslím, že dává prostor novinářům nebo redaktorům a editorům, jak si tu práci usnadňovat nebo jak najet na nějaké nové způsoby té práce. Ale osobně se úplně neobávám toho, že by přišli o práci v dohledné době. Jako nevíme, těžko říct... všechno tady to, o čem mluvím, je pro mě výhled jako roku, dvou. Myslím si, že ten růst je opravdu exponenciální, a tím pádem, že se opravdu může stát, že za dva roky se tady budeme bavit a už to bude úplně jinak. Ale to je věc, kterou musíme všichni přijmout, a i vyloženě s tím takhle pracujeme na té naší projektové skupině ohledně umělé inteligence v Českém rozhlasu, kdy říkáme: „Ano, tohle teď my víme a takhle se k tomu stavíme, ale zároveň víme, že to za půl roku může být úplně jinak.“

Ty jsi říkala ještě na začátku, že vždycky za to musí být odpovědný nějaký člověk. A tak mě u toho napadá, jestli by třeba dávalo smysl ustanovit nějakého jednoho odpovědného člověka v té redakci nebo o tom budete natolik edukovat třeba ty editory, aby každý z nich vlastně mohl být tímhle odpovědným s člověkem?

To je právě cílem toho interního dokumentu, který vzniká. Opravdu je potřeba, aby nejenom redaktori, ale aby všichni zaměstnanci Českého rozhlasu věděli, že tyhle nástroje tady existují a rozhlas nezakazuje s nimi pracovat, ale jsou tady nějaká pravidla. Teď to ještě nemá konkrétní formu, protože ten dokument vzniká, ale bavíme se o tom, jestli například nevzniknou nějaké checklisty, nebo my tomu říkáme semafor, kdy si ten jakýkoliv uživatel z rozhlasu řekne, že chce tady teď použít pro svojí práci dejme tomu... Já dávám za příklad ten ChatGPT, protože na tom se mi to dobře vysvětluje. Tak dejme tomu, že chce použít ChatGPT pro nějaký úkol, na kterém pracuje a bude mít k ruce od rozhlasu oficiálně vydaný nějaký checklist nebo semafor, který bude říkat: „OK, pohlídej si, aby si do toho hned nedával nějaká interní nebo citlivá data“. Pokud napíšu do ChatuGPT dejme tomu „jaké jsou trendy v doporučování obsahu v roce 2023 pro zpravodajské weby“, tak to je úplně v pořádku, protože to je nějaká moje brainstormingová diskuze s tím ChatemGPT. V okamžiku, kdy bych tam ale už napsala: „Mám takovouhle návštěvnost“ a nedej bože bych tam dávala třeba nějaké citlivé údaje těch uživatelů a podobně, tak to už je něco, u čeho mi ten semafor nebo ten checklist řekne: „Tohle ne, to tam přece nemůžeme dávat“. My nevíme, kam se ta data dostávají a nevíme, co se s nimi děje. Cílem té naší skupiny je kromě toho, že chceme sepsat nějakou interní směrnici nebo dokument, právě i to abychom přišli s těmi nástroji a metodami, jak ty uživatele tady v rozhlasu edukovat v tom, jak mohou ty nástroje umělé inteligence používat a jak ne. Zároveň je cílem je k tomu malinko popostrčit, protože rozhodně poslední, co bychom chtěli by bylo, aby si z toho někdo odnesl, že: „aha, tady je nějaká směrnice a abych něco nepokazil, tak to radši vůbec nebudu používat“. Naopak chceme edukovat a chceme, aby to ti uživatelé používali, protože vidíme ty výhody, které to má pro tu práci, ale má to vždy ten rámec, ve kterém to musí fungovat. Těžko říct, jestli jsme s tím jako na začátku, nebo jestli jsme s tím opoždění, protože si myslím, že různé firmy nebo různé mediální domy se k tomu staví různě. Někdo to ošetřuje dost striktně, někdo to neošetřuje vůbec. My tím, že jsme Český rozhlas, tak chceme být transparentní v tom, aby naši uživatelé a posluchači jasně věděli, jak s tím pracujeme a abychom mohli říct, jsme

inovativní, používáme to, ale přesně víme, jak to používáme a všichni naši zaměstnanci ví, jak tady s těmi nástroji pracovat.

Takže tomu dokumentu budou asi následovat nějaké workshopy nebo školení?

Zatím v tuhle chvíli nemáme přesnou představu, ale myslím si, že ten dokument by měl vzniknout ještě do konce tady toho roku a že ten příští rok bude přesně už takhle praktičtější, kdy budeme nějakým způsobem pracovat na tom, aby ti zaměstnanci věděli, co můžou, co nemůžou a k čemu jim to může být dobrý, k čemu jim to může pomoci.

Ještě, a to se vracím hodně na začátek, jak jsi mluvila o té home page a té stránce pro vás, tak teď tu homepage vždycky určuje editor nebo webeditor?

My na iRozhlasu vlastně nemáme webeditora, protože iRozhlas je celý webový. To znamená, že ta pozice se u nás jmenuje editor, ale je to v podstatě částečně web editor, který spravuje tu homepage, ale zároveň má ještě takovou tu klasickou roli editora, kdy opravdu řeší nebo zadává práci lidem v rámci třeba denní směny a řeší s redaktory obsah těch článků. A takhle je to kombinace obojího.

Budete teda čtenáře informovat o tom, že ty nástroje používáte, tak jakým způsobem? Jo, to už jsi vlastně zmínila, že teda stojíte za tím, aby to tak bylo...

Ano. Ono je to i zajímavé v tom smyslu, že tím, jak teď začaly přicházet pokročilejší nástroje, jako je právě třeba ta syntéza hlasu nebo Midjourney, tak jsme si uvědomili, že je potřeba opravdu tomu uživateli jasně indikovat, že je to výstup nebo práce umělé inteligence. A když to vztáhnou na to doporučování, tak dokud jsme měli ten boxík *Robot doporučuje* tak z mého pohledu to nebylo úplně označené přímočaře, že tady ten boxík je výsledek nějakého algoritmu. Protože *Robot doporučuje* je takový jako vtipný. Ale pokud se nad tím člověk nezamyslí, tak si nemusí uvědomit, že to je výsledek práce nějaké umělé inteligence. Proto i v okamžiku, kdy teď budeme ten boxík znova nasazovat, tak už se pravděpodobně nebude jmenovat *Robot doporučuje*, ale bude se jmenovat... Teď to není úplně jako ujasněné, ale já bojuji za to, aby se jmenoval *Doporučujeme* nebo něco takového. Ale bude u toho uvedené, že tohle je plocha, která je doporučovaná na základě informací, které o tom uživateli máme. Myslím si, že bude asi i dobré nějakým způsobem vysvětlit, co to je za informace, aby se ty lidi úplně nelekli, že o nich sbíráme nějaká data, ale jsou to vlastně úplně normální informace o procházení webu, která v okamžiku, kdy ten uživatel akceptuje cookies, tak se o něm prostě sbírají. Nedá se na základě toho ten uživatel identifikovat jako konkrétní uživatel, ale nám to přináší data o tom, jaké stránky navštívil, jaká témata čte, jaké autory čte. V tomhle smyslu určitě budeme chtít edukovat, aby se toho ti uživatelé nebáli, ale zároveň aby to začali chápat jako normální součást iRozhlasu. To je, co se týče boxíku *Doporučujeme* nebo *Robot doporučuje*. V okamžiku, kdy začneme pracovat víc třeba s tou stránkou *pro vás*, tak tam bych k tomu naopak já osobně chtěla přistupovat tak, že to doporučování je tam ta přidaná hodnota. Že díky tomu, že my o vás tady tato data máme, jsme schopni vám doporučit tady tyto články, které vás doufáme zajímají. Co tím chci říct, je to, že my jako rozhlas i u toho doporučování, které je už poměrně běžné na trhu a málokdo to označuje ve smyslu „tady toto vyhodnocuje nějaký algoritmus a na základě toho vám to doporučujeme“. Tam já vidím prostor a příležitost se k tomu postavit tak, že „ano, tohle my vám tady doporučujeme, nějak to zpracováváme ty vaše data a přináší vám to tenhle užitek“. A naopak to dát do popředí, aby ten uživatel byl opravdu srozuměný s tím, jak to funguje. Teď nechci slibovat nic konkrétního, ale dovedu si představit, že o tom napíšeme článek, kde napíšeme, jak to funguje a jak ta data vyhodnocujeme a jaké s tím děláme experimenty. Protože podle mě to

není na škodu, aby ten uživatel věděl, co to je přesně za data a že ho to nemůže nějak jako poškodit, protože mi na základě těch dat nemůžeme toho uživatele identifikovat, ale může mu to přinést něco dobrého v tom smyslu, že opravdu dostane ten obsah, který ho zajímá. Takže to je, co se týče označování doporučení. Co se týče označování, jak jsem už říkala, syntézy nebo Midjourney a takhle, tak tam je to velmi přímočaré. Tam opravdu u tlačítka *play* bude... my ještě nemáme tu konkrétní podobu, jo, ale určitě to bude označené nějakým piktogramem, který třeba po najetí myši, nebo i možná to tam bude přímo napsané... bude tam jasná informace, že „teď si chcete pustit audio, které bylo vytvořeno syntetickým hlasem“. Určitě to nawordujeme nějak líp jo, ale to sdělení je takovéhle. Protože já osobně si myslím, že to jsou i takové jako best practices, které se používají na webu, kde by měl uživatel vždycky vědět, co se stane v okamžiku, kdy někam klikne. Normálně v okamžiku, kdy třeba kliknu na titulek článku na homepage, tak už mám takový ten model jako uživatel, že: „ano, kliknu na to, otevře se mi stránka, kde bude detail toho článku, když kliknu na tlačítko *play*, tak vím, že se mi začne přehrávat třeba audio nebo video“. Podle toho, kde jsem a jak ten player vypadá, protože pokud je to velký obdélník s tlačítkem *play*, tak spíš předpokládám, že tam bude video. Pokud je to něco kompaktnějšího, tak spíš předpokládám, že tam bude audio. Je to samozřejmě doplněné i nějakým textem, ale uživatelsky tady tyto zkratky fungují. Už většinou víme, co máme očekávat, když něco uděláme. U těch nástrojů umělé inteligence, my jdeme vlastně o ten krok dál, protože ano ten uživatel ví, že když klikne na tlačítko *play*, tak se mu začne přehrávat nějaké audio, ale my chceme, aby už předtím, než na to tlačítko klikne věděl, že se mu začne přehrávat syntetické audio. Já si myslím, že jsme v tom opatrní, ale nejenom my, ale jako obecně se tady ta témata, jak to označovat, hodně řeší. Ale myslím si, že třeba za 2 roky už se to takhle nebude označovat, že už si na to všichni zvyknou a bude to normální. Ale teď jsme ve fázi, kdy uživatel, když kliknul na tlačítko *play*, tak očekává, že mu nějaký živý člověk bude něco říkat nebo mu bude hrát nějakou písničku. A v okamžiku, kdy tam není živý člověk, tak je podle mě férové a transparentní to tomu uživateli říct. A vůbec to nemění nic na kvalitě a hodnotě toho obsahu, který si pustí, ale je to prostě férové a myslím si, že bychom to měli dělat nezávisle na tom, jestli třeba potřebujeme vydělávat na reklamě, což jsou prostě komerční mediální domy, nebo jsme veřejnoprávní médium. My jsme teď na začátku nějaké fáze nebo éry, kdy se proměňují zažitá předpoklady, které ti uživatelé mají a je podle mě férové to těm uživatelům říct, aby nebyli překvapení.

Už jsi zmínila ten rozdíl mezi komerčním a veřejnoprávním. Tak jaké jsou podle tebe argumenty, proč to má opodstatnění v médiu veřejné služby? Proč ano, proč ne?

Já osobně to vidím tak, že veřejná služba tady musí být a je důležitá ze své podstaty. Aby byla uživatelsky zajímavá, tak musí držet krok s tím, co nabízí zbytek trhu. Což znamená, že pokud zbytek trhu nabízí doporučený obsah, nabízí dejme tomu audio verze článku, které jsou vytvářeny syntetickým hlasem... a my bychom to nedělali, protože bychom řekli: „Ne. My jako veřejná služba se toho bojíme,“ tak jediné, co tím způsobíme, bude, že od nás odradíme ty uživatele a takhle je přihrajeme k jiným médiím. Je to úplně pochopitelné, protože jako uživatel chci konzumovat ten obsah tak, jak je mi to nejpohodlnější. Například víme, že spousta lidí konzumuje audio obsah v autě po cestě do práce. V okamžiku, kdy bychom ten audio obsah neměli, a teď nemluvíme o rozhlasu, ale čistě o iRozhlasu, kde máme hlavně ty články... V okamžiku, kdy my víme, že uživatelé chtějí konzumovat audio obsah po cestě do práce v autě a my jim ho nenabídneme, tak oni si neřeknou: jé, no ježišmarja to je škoda, tak si ten iRozhlas přečtu třeba až dojeďu do práce. Ne, ten člověk si prostě pustí něco jiného v tom autě a my tam chceme být v tom autě s ním a je nad naše možnosti nahrávat

všechny články s živým člověkem. Ani by to nedávalo smysl z ekonomického hlediska nebo z hlediska smysluplnosti... kdyby tady měl někdo sedět ve studiu a načítat jeden článek za druhým, to ani není zajímavá práce, kterou by někdo třeba chtěl dělat. Když víme, že tady je nástroj syntézy hlasu, která nám umožní audio verze článků dostávat k těm uživatelům, tak kdybychom tímhle směrem nešli, tak tím jakoby zajistíme, že ti uživatelé si to najdou někde jinde. Proto z mého pohledu je důležité, aby média veřejné služby držela krok s tím trhem. Protože ten uživatel je prostě náročný v tom smyslu, že chce ten obsah tak, jak ho chce, tehdy, kdy ho chce a v takové formě, v jaké ho chce. A když my mu to nenabídneme, tak si ho najde někde jinde.

A ty hlavní důvody jsou teda finanční a kapacitní?

Já bych řekla, že naopak. Ten hlavní důvod z mého pohledu měl být o tom, že chceme tomu uživateli přinášet to, co on chce. Pokud máme identifikované, co ten uživatel chce, tak hledáme způsoby, jak se k tomu dostat a tam teprve vznikají ty finanční a kapacitní jako argumenty, proč to nedělat živým člověkem, ale proč to dělat třeba nějakou umělou inteligencí. Z mého pohledu je důležité prostě přinášet to, co ti uživatelé chtějí.

Mě napadá otázka na ty editory zase... Jaké by podle tebe měli mít ty nové kompetence? Protože chystám se ptát i jich, tak by mě zajímal ten druhý pohled, co byste od nich vy požadovali nebo jaké jsou ty důležité kompetence?

Určitě je to kritické myšlení, a to si myslím, že editoři obecně z podstaty své práce mají. To znamená, chápat, že pokud mi ChatGPT vyplivne nějaký text, že to nemusí nutně být fakticky nebo gramaticky správné a opravdu přistupovat k těm nástrojům tak, jako přistupuji třeba... tak část té editorské práce je o tom, že editor dostává články od redaktora, které potom kontroluje a vlastně říká, ano, tento článek se může publikovat. U tady té mezilidské spolupráce, kde editor dostane první článek od toho redaktora a řekne si, že to je dobrý nebo třeba mu řekne, ať upraví pár věcí, ale je s tím spokojený. A tak už si buduje nějaký vztah s tím redaktorem, a když dostane takových článků 10 od toho samého redaktora, které budou dobré, tak už potom trošku poleví, a ne v tom smyslu, že by už ten článek vůbec nečetl, ale už bude vědět, že tento redaktor mu vždycky dává dobré články, nejsou tam nějaké chyby, ať už obsahové nebo třeba gramatické. Tak když mi přijde 100. článek od tady toho redaktora, tak já už vím, že se zaměřím na to hlavní a po celé mé zkušenosti s tím redaktorem budu předpokládat, že to je v pohodě. A takhle vlastně funguje práce mezi editorem a redaktorem. Pro toho editora je důležité, aby si uvědomil, že pokud dostane 100 výsledků od ChatuGPT nebo od jakékoliv umělé inteligence, se kterými bude spokojený, tak tam nesmí vzniknout tenhle ten vztah. Že si po sto prvé řekne, že je to dobrý, vždycky se mu to osvědčilo, tak to pouští ven, protože se na to nedá spolehnout. Člověk si nemůže být jistý, že po sto prvé tam najednou nebude nějaký nesmysl. Takže tohle je třeba kompetence, kterou ten editor musí chápat. Zatímco na toho redaktora se postupně třeba může spolehnout, tak na nástroje umělé inteligence ne. Vždycky to je tak, že se to musí velmi důkladně podrobit analýze, ať už z toho obsahového hlediska nebo z hlediska stylistiky. Já jsem uvedla to kritické myšlení, ale je potřeba si uvědomit, že zatímco každý editor je inherentně schopný kriticky přemýšlet, tak na základě zkušenosti s tím redaktorem třeba může malinko polevit. Je důležité, aby chápali, že u té umělé inteligence nemůžou nikdy polevit. Obecně si myslím, že je potřeba, abychom je vzdělali trošku v tom, jak to vůbec funguje, z čeho se to učí, odkud to bere ta data... a to jsem si jistá, že někteří naši editoři absolutně neví, někteří jo, ale je to proto, že se o to třeba zajímají. Rozhodně to není tak, že by tady měli už nějaké jako školení od Českého rozhlasu a je to určitě bod, který budeme muset nějak řešit, aby byli proškolení.

My od nich nemůžeme očekávat, že jim takhle dáme nástroje a že jim třeba dáme i ten checklist, ale oni, když nebudou schopní si to vyhodnotit, nebudou chápat ty důsledky, tak po nich ani nemůžeme chtít, aby byly schopní to nějak zhodnotit. Takže nějaké obecné školení, jak tady ty modely fungují a z čeho se učí, jak s tím pracují, tak to je určitě potřeba, aby tady proběhlo a nebo aby probíhalo na nějaké jako pravidelné bázi. A je to něco, co absolutně v tuhle chvíli ještě nemáme nějak podchycené, ale máme to na nějakém seznamu věcí, které mají být výstupem té naší projektové skupiny.

Respondent Jan Cibulka, datum 17.10. 2023

Jak využíváte umělou inteligenci nebo její prvky v rámci iRozhlasu a potom v rámci datové redakce? A co budete využívat?

Já se zeptám, co považujete za umělou inteligenci pro rozsah vašeho výzkumu? Jsou to velké jazykové modely, generativní modely nebo všechno, co se tady pod to označení vejde?

Já jsem to nazvala prvky AI.

Když se zeptáte na prvky AI, tak já vám řeknu, ano, dlouhá léta, všichni, co píšeme články, tak hledáme na Googlu. Opravujeme si chyby ve Wordu. Někteří z nás mají telefon jako iPhone, to znamená, že tam fungují nějaké lokální modely třeba při opravě překlepů, ale to asi není to, co chcete slyšet, žejo?

Jo, tak já se teda zaměřuju na ty doporučovací algoritmy a doporučování v rámci mujRozhlasu a iRozhlasu.

Jakože doporučování obsahu?

Ano.

Oukej, s tím my nemáme lautr vůbec nic společného, to jde celé za Dalií. My jsme něco takového implementovali, ale jenom experimentálně před výrazně více jak 10 lety v Hospodářských novinách. Jako nějakou formu doporučování, ale to jsme dělali ne v rámci naší novinářské práce, ale v rámci nějaké vývojářské činnosti s kolegy. A jakože můžu říct, že ty věci i existují. To je asi tak to jako nejlepší, co o nich umím říct, nebo jako... Umět nějakým způsobem algoritmičtě doporučit obsah není nic extrémně objevného. Koneckonců používá to každý e-shop dneska. Samozřejmě je s velkým otazníkem, jestli to dané médium produkuje dostatek relevantního obsahu, aby vůbec něco takového bylo nutné implementovat.

Jednak je tu nějaká obava z uzavírání do těch filtrovacích bublin. Ale to si myslím, že je ta je do značné míry falešná, protože drtivá většina zpravodajských médií neprodukuje takové množství obsahu, abyste v nich reálně ty filtrovací bubliny mohla vytvářet. Snad s výjimkou toho, že prostě lidi, co čtou jenom sport, tak budou číst jenom sport. Ale jinak toho obsahu prostě není tolik. iRozhlas vydá nějak do 60 článků denně. V tom je nějaká otázka rekomandace s velkým otazníkem. Samozřejmě máme tady něco trochu jiného, ale to máme jako Český rozhlas, ale s tím já nemám nic extra společného, to jenom vím, že to funguje. Máme, myslím, nějakou rekomandaci obsahu na serveru mujRozhlas, kde jsou audio pořady a tam to do nějaké míry smysl dává, protože prostě jsou lidi, co vyhledávají kriminálky, tak jim nabídneme další kriminálky. Jsou lidi, co chtějí poslouchat četbu na pokračování, tak dostanou četbu. Ale to není technologicky nějak extrémně komplexní. Asi bych se i zdráhal pro ty algoritmy, které tady to zprostředkovávají, používat označení AI, rozhodně v tom současném chápání je to spíš nějaký jako algoritmus strojového učení v lepším případě. V horším případě je to prostě série, nějaký jako rozhodovací strom na základě třeba klíčových slov nebo nějakých rubrik.

A co vy osobně používáte při práci?

Pokud se máme bavit o tom, co my teď používáme v datovém oddělení, nebo obecně na iRozhlase, tak máme nasazené experimentálně pro naše redaktory nějaké generativní modely. Jedním z nich je ChatGPT od OpenAI, který máme v několika podobách nasazený, ale v zásadě je to upravená konzole do ChatGPT 3.5 a 4, která je nějakým způsobem jako

předpromptovaná na otázky, co by redaktori mohli používat. To je jakoby teorie. My jsme to nasadili hodně jednoduše technicky, abychom se drželi nízko při zemi. Zkušenost je taková, že pro naše redaktory to valného přínosu nemá, protože prostě ty schopnosti, řeknu teda těch jazykových modelů, jsou dneska hodně basic a na práci ve zpravodajství se v tomhle tom směru zatím příliš nehodí.

Na co to používáme jako dost je kontrola gramatiky. Konkrétně si to pochvaluje kolegyně, která hodně času v životě žila v zahraničí, to znamená, že její zručnost s českou gramatikou není taková. Tady ty generativní nástroje jsou efektivnější nežli nějaký Microsoft Word nebo něco na ten způsob, protože jsou schopná chytat různá vyšínutí z větné vazby a věci, které jsou v takové té základní formální gramatice v pořádku, ale následně jsou nesmyslné nebo vybočují z nějakého úzu. Dále používáme Whisper od OpenAI, což je open source model pro přepis zvuku nebo mluveného slova do textu. Konkrétně Whisper je pro nás opravdu veliký kvalitativní skok, protože my jsme nějaké komerční nástroje v minulosti používali na přepis z mluveného slova do textu, ale ta kvalita byla extrémně nízká (tady těch modelů). Navíc byly výpočetně náročné a dost pomalé. Whisper všechno tady to odbourává. To znamená, je open source, je zadarmo, je velmi rychlý, pokud je nasazený správným způsobem, a umí velmi dobře česky i na nízko kvalitních nahrávkách. Typicky nahrávky z ulice a podobně.

Jsou ještě nějaké další takové nástroje, které že by vám v budoucnu nějak pomohly usnadnit práci?

Těch generativních jazykových modelů je víc. My máme ještě zkušenost s Claudem, což je produkt od společnosti Anthropic. Má nějaké výhody, nějaké nevýhody před těmi produkty od OpenAI, ale ty výsledky v tom závěru, k čemu se dají použít, nedají použít a jakým způsobem se dají nasadit na novinářskou práci, jsou vlastně dost podobné. To, k čemu my to v budoucnu budeme velmi pravděpodobně používat, protože s tím úspěšně experimentujeme, je extrakce informací z velkého množství textů. To znamená třeba nějaké jako sdružování, řeknu rozhodnutí soudů nebo nějaké jako štítkování článků pro metody interního vyhledávání. Věci tady toho typu, řeknu spíš analytické, spíš zaměřené dovnitř. Ne že bychom tu produkci, kterou ten model jakoby vyvrhne, potom předložili čtenářům, protože na to je to často fakticky mimo a chybí tomu hluboká znalost problematiky. Třeba teď jsme testovali, jestli jsme schopni ten model nějakým způsobem doučit, aby lépe reflektoval to, jakým způsobem na iRozhlase děláme titulky k článkům a byl schopný takové titulky generovat. My jsme schopni vlastně ten model doučit způsobem, že se bude držet těch formálních náležitostí. Ale co se týče obsahové stránky, tak ta bude výrazně míň kvalitní než to, co dělají naši editoři. A vzhledem k tomu, že ten titulek do značné míry rozhoduje o tom, jestli to ti čtenáři budou číst nebo nebudou, tak tam chceme tu kvalitu spíš jako zvýšit, ne snížit.

A jsou nějaké specifické nástroje třeba pro tu datovou redakci?

Neřekl bych úplně specifické prvky strojového učení. Určitě používá Google Pinpoint, což je nástroj pro extrakci strukturovaných informací z nestruturovaných textů z dokumentů. Je to jedna z těch věcí, kterou prostě testujeme nebo užíváme.

Co podle vás využívání takových nástrojů přinese čtenářům?

Myslím si, že to není jako nejbližší výsledek, který uvidíme v akci, protože my už to tak testujeme. My se jako novináři v nějakém našem omezeném čase budeme moct zaměřit na hlubší analýzy větších objemů textů. V minulosti, jsem si nemohl říct, že ručně přečtu stovky

rozsudků k nějakému jednomu tématu, protože jsem na to prostě neměl čas, ani když se mi ty rozsudky podařilo získat. A dělali jsme to jenom v případě opravdu jako důležitých článků, abychom na to mohli vyčlenit tu lidskou činnost. Dneska tohle budeme moct dělat třeba u širšího rozpětí témat, na kterých bychom normálně ty lidi vyčlenit nemohli, ale díky tomu, že máme k dispozici tady ten nástroj, tak i s tím, že je do nějaký míry chybový, tak pro nás je to pořád prostě lepší než nic. Čtenáři v tom budou benefitovat v tom smyslu, že my budeme schopni lépe pokrýt některé specifické druhy témat rychleji, lépe a vůbec se jim třeba budeme věnovat, protože někdy bychom to s ohledem na lidskou náročnost úplně vynechali, protože na to nemáme lidi.

Tak to jste zmínil vlastně nějaké výhody. Přinese to třeba i etické výzvy?

Etické výzvy to určitě přináší, protože spousta lidí zejména teda v managementu, a tím nechci říct jako v rozhlasovém managementu, ale obecně v managementu médií po celém světě, tak je hodně jako lákaná, protože média nejsou v dobré finanční situaci, tak je lákaná do toho: „Hele budeme to používat ke generování textu, který vypadá vlastně velmi uvěřitelně, je napsaný gramaticky správně a budeme ho teda nabízet našim čtenářům a čtenářkám“. Nic takového v rozhlase neplánujeme, protože víme, že ty věci prostě nejsou fakticky správně, a ta kvalita je taky k debatě. Tímhle způsobem můžete produkovat pouze obsah nízké kvality a často velmi nepřesný, až nepravdivý.

Podívejte se, jak vypadají popisky zboží na nejmenovaném e-shopu. Ty věci jsou generované, jsou fakticky správné a jsou informačně naprosto nehodnotné. Vy se z popisku zboží jakožto uživatel vůbec nic nedozvíte. Zejména ve chvíli, kdy jdete koupit komponentu do počítače, kde jsou jako zásadní ty detaily a nepotřebujete vědět, že „tento disk je opravdu velký a opravdu rychlý, což z něj dělá ideální disk do vašeho počítače“. To je typická produkce, kterou vám bude schopný vytvořit nějaký velký jazykový model, který bude držet ty formální znaky takového popisku a bude držet formální znaky toho použitého jazyka, ale nebude schopný uplatnit nějakou vnitřní znalost problematiky.

Jak využívání prvků umělé inteligence promění přímo vaši práci anebo práci editorů na iRozhlase?

Já si myslím, že práci editorů na iRozhlase to nijak zvlášť nepromění. V tom smyslu že jasně, můžeme to použít pro opravu gramatiky a stylu. Tam to smysl může dávat v některých případech. Nikoliv to ale nemůžeme použít ke kontrole faktické správnosti textu, protože samo je to schopné halucinovat, natož aby to ještě kontrolovalo správnost nějakého jiného textu. To samé různá hodnocení vyváženosti a podobné další věci. Naše zkušenost, kterou jsme s tím udělali, tak je vlastně velmi špatná. Věci, u které lidský editor jasně identifikuje nějaký přítomný bias, tak tady ta věc je prakticky schopná vždycky označit prakticky za vyváženou, protože není schopná odlišit kontext, ve kterém daný text vzniká, a jednotlivé nuance.

To znamená, že jako práci editorů to asi nějak extrémně nepromění, ani si nemyslím, že by to nějak extrémně proměnilo práci redaktorů s tou výjimkou, že jak říkám, díky jazykovým modelům, které máme dnes k dispozici, budeme schopni dělat efektivnější textový mining ve větším objemu informací. Až se nám prostě podaří definitivně přesvědčit Ministerstvo spravedlnosti ke zveřejnění rozsudků v České republice třeba v trestních kauzách, tak budeme schopni z tady toho velkého množství textu poměrně rychle dělat analýzy toho, jakým způsobem se rozhoduje o nějaké problematice.

A jsou podle vás editoři nebo novináři připravení na to pracovat s tím? Kde vy jste se

třeba naučil s tím pracovat?

To znamená, říct, jestli jsou novináři připraveni na práci s novou technologií v České republice. Tak velmi pravděpodobně většina z nich není, ale to neznamená, že se to prostě nenaučí. Jenom v tomto směru bude potřeba nějakého jako dalšího vzdělávání. Jsou tady nějaké snahy ze zahraničí, to znamená typicky třeba Google spolu FT dělá školení pro novináře na používání generativní umělé inteligence. Není to, kdovíjak dobré a samozřejmě je to trošku zaměřené na prodej konkrétních produktů, protože to dělá konkrétní technologická firma. Víím, že se ke školení pro novináře chystá třeba Nadační fond nezávislé žurnalistiky, který už dneska novináře školí.

Já to třeba ťukám na některých svých přednáškách, co dělám pro studenty žurnalistiky na Karlovce. To znamená, že formou nějakého vzdělávání se na to novináři nějakým způsobem adaptují. On taky už do značné míry opadá ten *hype*, kdy i Česká média velmi nekriticky psala: „Tohle nahradí statisíce lidí ve výrobě, v provozu, v tvorbě textů, tohle nahradí právníky, tohle nahradí novináře“. Tím, jak lidi začínají už mít reálnou dlouhodobější zkušenost s tou technologií, tak vidí, že se nic takového neodehrává, ale že zase naopak objevují způsoby, jakými se ten nástroj dá použít. Ale jak to obvykle těchto nástrojů bývá, nic z toho není jako: „Tohle změnilo lidskou civilizaci.“ Je to: „Tohle je užitečný nástroj, který může řešit některé otázky, které jsme třeba do teď vyřešit uměli, ale za daleko větších finančních a časových nákladů.“

A v rámci třeba iRozhlasu, co by editorům nebo redaktorům nebo i vám v tomhle směru pomohlo? Jsou to nějaké ty workshopy?

Přemýšlím... Samozřejmě, vždycky je dobré vidět, co někdo vyzkoušel a jaké to mělo důsledky. Proto my třeba AI testujeme u nás na iRozhlase formou takových maličkých experimentů. Protože prostě jsme si někde přečetli: „Hele ta věc by mohla dělat věci X,Y a Z. Následně tedy jdeme, vymyslíme vlastně maličký experiment, který provedeme, následně ho vyhodnotíme a řekneme: „Ano, na to se to použít dá,“ nebo „ne, na to se to třeba spíš použít nedá,“ nebo třeba „teď to na to použít nejde, ale do budoucna možná najdeme cestu, jak to využít.“ Čím víc nějakých subjektů by tohle dělalo a ty znalosti sdíleli. Kdyby tohle někdo vyzkoušel před námi a ukázal nám, jak to dopadlo, tak to nemusíme zkoušet my. Ale nejsme si vědomi toho, že by nějaká taková extra výměna znalostí v České republice probíhala, proto se to snažíme jak já, tak i kolega Michal Kašpárek zprostředkovat třeba jiným redakcím. Když si o to řeknou, tak jsme ochotni jim udělat nějaké školení, kde představíme naše zkušenosti s konkrétními pokusy o nasazení té technologie.

Když bych se zaměřila teď na ty doporučovací algoritmy. S Dalíi jsme se bavily o *Robot doporučuje* nebo třeba o nějaké možnosti mít do budoucna i personalizovanou homepage iRozhlasu. Tak by mě zajímal váš názor na to.

Tady musím vytknout před závorku, že já tohleto nedělám a nenasazuji na iRozhlase. Ta technologie nejde za mnou, je to externí dodavatel a věc Dalie, takže s tím jako nemám nic extra společného. Nicméně v obecné rovině, já jsem spíš jako skeptický k té ideové části toho. Protože jak jsem říkal. Já si nemyslím, že máme takové množství obsahu, a tak široce rozkročeného, že by čtenáři potřebovali doporučovací algoritmus, aby se v něm zorientovali. Naše texty obvykle velmi rychle zastarávají. Ve zpravodajství není nic staršího než včerejší zpráva. To znamená, že jsem i hodně skeptický k tomu, že budeme schopni formou doporučování oživovat starší články. Zkrátka protože doporučovací algoritmus bude schopný určit: „Ano, tento článek se velmi týká toho tématu, ke kterému ho doporučuje.“

Nicméně ten článek bude rekapitulovat dění, které už je překonané nějakým dalším zpravodajským vývojem. To znamená, že tomu čtenáři spíš bude předkládat informace, že on si přečte článek ze dneška a následně mu tedy bude doporučen článek starý třeba měsíc, ale reálně se dostane k tomu, co už mu shrnuje ten dnešní článek mínus ten vývoj v dnešním článku. To znamená ta hodnota bude extrémně nízká. Fakt si já osobně nemyslím, že máme tak strašně široce rozkročený obsah, aby tento mechanismus našel velké uplatnění. S výjimkou toho, že prostě jsou lidi, kteří čtou jenom sport a nic jiného je nezajímá, a pak jsou zase lidi, kteří naopak jako velmi intenzivně nechtějí číst sport, ale to se dá prostě vyřešit tím, že si uděláte na webu rubriku sport a jsou lidi, co tam chodí, jsou lidi, co tam nechodí.

Jsou tedy nějaké inovace v téhle oblasti, které byste přivítali, nebo ani ne?

Záleží, no já osobně v tomhle jsem spíš dost konzervativní. Za prvé z těch důvodů, že tady ty algoritmy pro svůj provoz potřebují informace o předchozím chování uživatele na tom webu a v případě tak široce rozkročeného webu jako je iRozhlas, tak ty informace o těch přečtených článcích nemusí být nutně extrémně vypovídající, protože tam člověk prostě najde denní zpravodajství. Ale ve chvíli, kdy takovouhle technologii začnete nasazovat na újeji zaměřeném webu třeba investigativním, něco jako je Hlídací pes nebo Investigace.cz, tak ty informace o tom, co tam ten člověk čte a že tam vůbec chodí, tak jsou ve vztahu k němu jako velmi identifikující. Stejně tak pokud se budeme bavit třeba o webech, které se zaměřují nebo specializují na poskytování obsahu pro nějakou sexuální menšinu nebo pro lidi s nějakým onemocněním, tak informace o tom jejich procházení je vlastně poměrně citlivá, a proto já jsem dost opatrný v tom tady ty informace shromažďovat, a následně je používat k nějakým bohulibým cílům. Protože ve chvíli, kdy jsou shromážděné, tak můžou utéct a můžou případné subjekty těch informací do nějaké míry poškodit. Druhá věc, která s tím úzce souvisí a vyzkoušel si to poměrně nedávno Netflix: Ve chvíli, kdy začnete profilovat lidi na doporučování obsahu, velmi snadno se vám může stát, že aniž byste to zamýšlela, skončíte s kategoriemi, které spadají do kategorie těch zvláštních kategorií osobních údajů. To znamená informace o politickém směřování, náboženském přesvědčení, sexuálních preferencích a podobně. V tu chvíli jsou to informace, které už vy jako neuhájíte jako oprávněný zájem podle GDPR a vůbec byste je jakožto zpravodajský web zpracovávat neměla. A je velmi těžké se tady tomu biasu vyhnout. To znamená, že i na to je potřeba dávat si pozor, a i proto já jsem skeptický k nástrojům automatického doporučování. Plus samozřejmě musíte mít od těch lidí vědomý souhlas. To znamená nějakým způsobem je k tomu profilováním vyzvat skrze cookie lištu, což je další „otravný prvek dnešního internetu“. A tím, že ji spousta lidí ignoruje, tak vám ten souhlas vlastně nedá. Samozřejmě jsem zvědav, jak dopadne ten experiment, který dělá Dalia, ale viděl bych tam prostě tady ty překážky, které nesouvisejí přímo s tou technologií jako takovou, která, jak víme, může existovat a může dělat to, co tvrdí, že dělá. Ale neřekl bych, že ty výsledky budou jako oslnivé.

Ještě pro upřesnění, co se stalo na tom Netflixu?

Tam šlo o to, že na Netflixu ve chvíli, kdy si tam něco pustíte, tak vám pak ukazuje další podobný obsah. To znamená, že když tam pojedete nějaký teen romance, uvidíte víc teen romancí. Netflix zjistil, že má nějaké množství seriálů nebo obsahu obecně, který je zaměřený primárně na queer lidi. A ti lidi ho samozřejmě konzumují, ale v tu chvíli se dostávají do nějaké kategorie, která ale jakoby identifikuje jejich sexuální orientaci. V tu chvíli už je to zvláštní kategorie osobních údajů, u kterých podle GDPR nejde to zpracování překlenout jenom souhlasem. Musí tam být oprávněný zájem, a ten v rámci provozování

serveru s videi prostě právně nepostavíte. To znamená, že se dostanete do rozporu GDPR a hrozí vám pokuta. A tady to riziko obsahují prostě všechny zpravodajské weby, protože neuhádáte, že v rámci oprávněného zájmu na nějaké personalizaci obsahu máte informaci o sexuální orientaci vašich čtenářů nebo informace o zdravotním stavu. Můžete mít rubriku *zdraví a nemoci* a lidi, kteří mají rakovinu, si tam budou číst věci o rakovině, což je logické, a tím, že vám spadnou prostě do kategorie *rakovina slinivky*, tak máte najednou informaci o jejich zdravotním stavu, kterou byste prostě neměla zpracovávat vlastně za žádný účelem jako zpravodajský web, a proto je potřeba na tohle být hodně opatrný. Samozřejmě dá se to řešit technicky tím, že ty kategorie začnete rozostřovat. Ale zase v tu chvíli už nemáte na běžném českém zpravodajském webu takové množství obsahu, aby vám teda ty rozostřené kategorie něco smysluplného řekly. Zase s výjimkou jako sport/nesport, možná lifestyle, pokud ho na tom webu máte. My ho třeba na iRozhlase nemáme, takže tam jako není co řešit.

A v té současné chvíli, teď by se to dalo zjistit podle cookies? Nebo v budoucnu třeba by ten člověk musel mít založený nějaký svůj profil?

To jsou spojené nádoby. Vy potřebujete člověka nějak identifikovat v čase, což nejjednodušším způsobem uděláte prostřednictvím nějakého účtu, který pokaždé, když se ten uživatel přihlásí, tak mu to na ten jeho prohlížeč šoupne cookinu, která identifikuje. Takže víte, že je přihlášený a následně s tím, teda nějakým způsobem pracujete.

Tohle mimochodem přináší zase spoustu, nechci říct nedomyšlených důsledků, ale jako technické náročnosti. Poměrně rychle vám roste ten technický overhead. Najednou musíte spravovat informace o těch uživatelích. Musíte jim dávat možnost ty účty vytvářet, rušit je, měnit v nich informace, s ohledem na GDPR jim musíte dát možnost si nějakým způsobem vygenerovat a stáhnout všechna data, které s účtem souvisí, ten účet zrušit a smazat. Přináší vám to poměrně velké množství nějaké práce navíc a je otázka, čím by to vlastně mělo být vyvážené. Jestli to má být ta personalizace, tak je zase s velkým otazníkem, jestli ta personalizace bude vykazovat natolik zajímavé důsledky, že pro ty uživatele bude natolik zajímavá, že vám to za všechny tady ty technologické problémy vlastně stojí.

Ještě se zeptám, s ohledem na to, že to je zpravodajská služba veřejnoprávního média, tak jestli podle vás, už jste to vlastně částečně zodpověděl, ale jestli to je oprávněné vlastně takovou službu mít? A proč ano a proč ne?

Já si myslím, že nám nic právně nebrání to vyvinout a používat. V zemích jako Austrálie nebo velká Británie je zakotvený i legislativně koncept, a mimochodem v novele zákona o Českém rozhlase a České televizi je to už navrženo i v Česku, že veřejnoprávní média mohou sloužit jako nějaký inkubátor inovací právě pro soukromý sektor.

Dám příklad. Vyzkoušíme nějaký doporučovací algoritmus. Zjistíme, že opravdu nějakým způsobem zázračně funguje, je to hrozně dobrý a čtenáři to milují. Tak v tu chvíli je vlastně, řekl bych, naší společenskou povinností, nejenom, že tedy to máme a zkusili jsme to, ale zároveň otevřeně mluvit o tom, že to tak funguje a publikovat ty výsledky. To znamená, že ten soukromý sektor to vlastně velmi jednoduše může převzít a nemusel vynaložit náklady na to testování.

Naopak ve chvíli, kdy my v rámci testování zjistíme, že to za nic nestojí a že to nepřináší ty výhody, tak je rovněž, řekl bych naší společenskou povinností tady tu informaci publikovat i s nějakým popisem toho, co jsme udělali, aby potom ta soukromá média si mohla odškrtnout: „Dobře. Na tohle jsme nemuseli vynaložit peníze. Víme, že to za nic nestojí, nemusíme to dělat.“ Veřejnoprávní média pro ten soukromý sektor vlastně můžou být

přínosem jako inkubátor nějakých inovací, aniž by mu byly konkurencí. Protože my mu nejsme konkurencí v tom obchodním smyslu, protože my nenabízíme online reklamní prostor, ani nenabízíme předplatné a ty naše služby si posluchači a čtenáři předplácí jinak prostřednictvím koncesionářského poplatku.

Z pohledu uživatele, je to pro něj výhoda, nebo nějaká přednost toho média, kterou by podle vás třeba ocenil, anebo si naopak řekne, že je v pořádku, že rozhlas doporučovací technologie nevyužívá?

Cílem toho doporučování je zvýšit nějaký uživatelský komfort. Jako tím deklarovaným cílem. Je otázka, zdali se toho podaří dosáhnout. Já na tohle jako nemám odpověď. Říkal jsem, že jsem z nějakých důvodů, které jsem zmínil, k tomu trochu skeptický, ale myslím si, že tady to je možné rozhodnout až ve chvíli, kdy zjistíme, jestli to k něčemu je nebo není. Samozřejmě ve chvíli, kdyby se to z nějakého důvodu ukázalo jako extrémně užitečné pro uživatele našeho webu, tak proč bychom to neměli mít a neměli jim poskytnout lepší službu. Ale jak říkám, nemyslím si, že to množství obsahu, které produkujeme my, je tak široké, že by to byl jediný nástroj, jak toho uživatele, na tom webu zorientovat. Netvrdím, že to nemůže pro někoho, kdo má širší tu produkci, být smysluplnější. Řeknu nějaké velké vydavatelství jako CNC, které dělá celou šíři titulů, tak ve chvíli, kdy by zavedla personalizaci obsahu napříč všemi těmi tituly, tak třeba tam už by ten objem byl natolik široký, že by to nějaké reálné benefity mělo.

Další otázka, je doporučování v oblasti zvuku. Tam na to nejsem nějak vůbec kovanější, ale tam si dovedu představit, že si prostě lidé poslechnou jeden pořad a řeknou: „Doporuč mi pořad podobného typu nebo díl pořadu podobného typu.“ Máme tam nějaké rozhovorové pořady, třeba Bány Tachecí, kterým prošly stovky lidí a teď prostě někomu se ten pořad líbí, ale chce si poslechnout díly, kde vystupuje vědec, a ne třeba politik, nebo ne třeba komentátor. A pokud budeme mít nástroj, který bude schopný tady to přefiltrovat, tak si dovedu představit, že si to někdo jako může tady tím způsobem přefiltrovat a poslechnout. Ale to už odbočujeme od toho iRozhlasu, tady to je jiný obsah produkovaný rozhlasem.

A pokud by teda iRozhlas využíval nějaký prvky AI nebo to doporučování, tak...

No já jsem jako trošku na jehlách z toho, že vy tam mícháte AI a doporučování. Doporučovací algoritmy, jako já se i bojím tomu říkat strojové učení. To je prostě statistika, jako nějaká vlastně poměrně banální. Jasně, můžeme si asi říct, že na to nalepíme tu nálepku AI, ale je to spíš marketingová nálepka. Reálně doporučovací algoritmy jsou prostě jako série 4 „ifů“, když budu jako hodně zlej. Takže bych byl opravdu opatrný tomu říkat i umělá inteligence.

Tak zůstaňme jenom u těch doporučovacích algoritmů. Takže kdyby se používaly na iRozhlasu, tak byste to nějak označovali pro ty čtenáře? Aby si ten čtenář byl vědomý toho, že mu to doporučil algoritmus?

My už to tam teď máme. Proto jsme tam měli to *Robot doporučuje*. To je jakoby implicitní označení toho, že je to doporučované strojem. I do budoucna by to v nějaké formě bylo. Můžeme si představit sekci, která se prostě bude jako jmenovat *doporučeno* nebo něco na ten způsob. Ještě se s tím pojí jedna věc. Kam ty doporučení umístíte? Protože když se podíváte na náš web, když si tady něco záměrně rozkliknu... vy si otevřete nějaký náš článek, je tam ten text a po pravé straně máte vlastně rubriku „zprávy, které jste ještě nečetli“. A když ten člověk dočte až ke konci, tak jsou tam ještě nějaké další. „Mohlo by vás zajímat“, tam je další čtveřice a pak je tam přehled nejdůležitějších zpráv z domova. A takhle to jde

dál a dál. OK, někde tady by mohlo být to doporučení třeba právě vpravo v té sekci „zprávy, které jste ještě nečetli“. Ale když si ten samý článek zobrazíte na mobilním telefonu, tak najednou zjistíte, že kromě toho textu článku, už moc nemáte místo, kam byste to doporučení mohla umístit. Respektive jediné místo, kam ho můžete dát, je po přečtení celého toho článku: *související článek a zprávy, které jste nečetli*, to se vlastně zařadí za ten text. A teď jaká je realita? Církou 85 % všech našich čtenářů přijde z mobilního telefonu. Takže ono není moc prostoru, kde to doporučování umístit, protože by to předpokládalo, že budou muset dočíst ten článek až nakonec, což zase z jiných měření víme, že lidi dělají absolutně výjimečně. Čtení článku vypadá jako takové velké *písmeno F*. To znamená, že lidi si přečtou celý titulek, většinu perexu, a pak projedou přes některé odstavce a musí do toho tématu být jó zažraní, aby to opravdu dočetli až úplně do konce. To znamená, že tady přibývá i ta technická otázka, kde by se jim vlastně to doporučování realizovalo. Tam spíš dává smysl... ten doporučovací algoritmus může fungovat dvěma směry. On totiž nutně nemusí pracovat na venek, ale on může identifikovat podobné články, které následně zobrazuje autorovi při psaní textu a ten autor následně potom ty ostatní články, které souvisí s tím jeho textem pod nějakým vlastním zralým uvážením v tom článku odkáže a řekne, jak iRozhlas psal už v minulosti: „Premiér řekl tohle a tamto.“ A tam už to smysl dává, ale to není nějaký algoritmus, který jako běží na pozadí těm čtenářům a sbírá o nich informace. Ale to je algoritmus, který funguje interně, nebo by mohl fungovat interně v našem redakčním systému a pomáhat vlastně autorům ty články lépe propojovat v tom tématu. Protože ty lidi se na tom webu střídají. Je ranní směna, odpolední směna. Někdo má dovolenou. To znamená, že někdo o tom tématu psal minulý týden a dneska o tom tématu může psát někdo úplně jiný, který prostě nemá znalosti toho člověka z minulého týdne, protože měl zrovna dovolenou. A v tu chvíli by bylo dobré tomu člověku, který píše ten článek říct“ „Hele, ale my už jsme o tom psali v minulosti a můžeš z toho vyjít a můžeš na tom stavět a můžeš na to odkázat.“ A tam my mimochodem tohle užití přijde daleko užitečnější v tom smyslu, že odpadá celé celé ten bohlav s nějakými uživatelskými účty, s GDPR a profilováním lidí, a naopak tohle jde na ruku těm autorům, kterým to pomáhá efektivněji pracovat, takže jsou prostě rychlejší, jejich články mají širší kontext a jsou napsané líp.

Když byste to shrnul, tak jste skeptický vůči doporučování kvůli GDPR a kvůli těm citlivým údajům, ale takhle v redakci jako v redakčním systému, by to použití bylo v pořádku?

I kvůli technologickým překážkám. Ty bych asi jako napsal na první místo. Samozřejmě, bylo by hezké říct, že nejvíc se staráme o GDPR, ale vzhledem k tomu, že v České republice má jeho vymáhání na starosti Úřad pro ochranu osobních údajů, který nemá absolutně nikoho s technickou znalostí, a to vymáhání prostě neprovádí, tak jakoby GDPR obecně, velmi cynicky řečeno, v České republice není moc velký strašák, protože to nikdo nedělá. Řekl bych, že ty technické překážky a technická náročnost nasadit nějaké „profilování uživatelů“, tak tam si myslím, že výhody rozhodně nepřevažují. Ale to nasazení dovnitř, aby my jsme byli schopni psát lépe, to si myslím, že tam to smysl dává, všechny tady ty problémy odpadají. To mi přijde jako smysluplné užití. A to je i samozřejmě něco, co my máme v plánu vyzkoušet v budoucnu. Takže to není jenom jako něco, co si tady vymyslím, ale něco, co bychom do budoucna nějakým způsobem chtěli používat.

A jaké jsou tedy všechny ty technické důvody?

Ty důvody, co jsem zmiňoval. Musíte vytvořit a spravovat databázi těch uživatelů, musíte hlídat jejich osobní údaje, musíte se starat o takový ten běžný provoz jako zakládání účtu,

rušení účtu, ověřování emailových adres a měnění informací. Například ten člověk třeba změnil email a chce ho vyměnit v tom účtu nebo chce výpis všech dat, která o něm v souvislosti s tím účtem uchováváte nebo ten účet chce zrušit a chce všechna ta data smazat. Všechno tady to musíte vlastně řešit už jenom tím, že tam máte nějaké uživatelské účty. Lidi zapomínají hesla a musíte jim dát způsob, jak si ta hesla obnovit. Musíte jim dát možnost se na vás obrátit ve chvíli, kdy jim někdo to heslo ukradne do toho účtu se jim nabourá a tak dále.

Respondentka Kateřina Cibulka, datum 23.10. 2023

Používáte nástroje jako třeba chat GPT, Whisper nebo Midjourney při vaší práci?

Při své práci úplně ne. Nebo ne v tuctle chvíli. Mluví se o tom, že by byla zvětšena kapacita toho, jak využíváme umělou inteligenci. Že by to třeba editovalo některé chyby. Například při přebírání agenturních zpráv jsou tam některé opakované jevy, které aktuálně vždycky musí opravit někdo na směně a musí poznat, co je špatně. A protože se ty jevy stále opakují, mluvíme o tom, aby se to zlepšilo. Mluví se taky o tom, že po hackerském útoku v létě na archiv Českého rozhlasu, by byla část obrázků nahrazená Midjourney, tedy ilustračními obrázky. Ale to všechno je vlastně ve fázi debat nebo plánů. Zatím nic z toho úplně nepoužíváme. Já jsem editorka a můj úkol je z větší části dívat se na to, co napsaly regiony Českého rozhlasu nebo jiné než zpravodajské stanice, vybírat z toho ten nejvíc zpravodajský obsah a upravovat ho tak, aby měl tu formu pro iRozhlas. Takže máme nějaké nástroje, které by nám mohly pomoci s našeptáváním, ale ten našeptávač, který je v Drupalu je středně spolehlivý, takže ho vůbec nepoužívám. Jediné, co v tuctle chvíli používám, je nějaká automatická kontrola překlepů.

A o jaký konkrétně nástroj jde?

Máme na to tlačítko přímo v našem Drupalu, v našem systému, a to je schopné odhalit překlepy a případně nabídnout to, že je rovnou přepíše, tak to používáme nejčastěji. Potom si věci házím ještě do Wordu, který dokáže odhalit chybějící čárky. Takže to jsou nástroje, které by se nám hodily, ale mluvíme právě o tom, že bychom zavedli něco, co nám například změní malé s ve slově sněmovna na velké S, což je teď něco, co vyžaduje lidskou práci. My to požadujeme po našich směnařích, aby se tohle naučili a zorientovali se v tom. A je to neuvěřitelně otravný a neustále se opakujeme, či pálíme na tom čas. Manžel Honza Cibulka o tom mluví, že by to chtěl zavést, takže až to zavede, tak se všem hrozně uleví. Ale v tuctle chvíli to moc nepoužíváme.

Jsou ještě nějaké další věci, které by vám pomohly? Napadá mě třeba přepis?

Vlastně ano, přepisovač máme. Právě díky Honzovi Cibulkovi máme k dispozici dva přepisovače. Jeden je ten, který už máme zavedený v našem Drupal systému, ten myslím dodává firma SpeechTech a ten je pomalý a trochu méně spolehlivý, ale jeho výhodou je, že přepisuje slovo od slova vysílání všech stanic asi týden dozadu, máme možnost i hodinu po hodině, což je skvělé. Když si potřebujeme stáhnout a přepsat rozhovor nějakého experta, který běžel ve 14:30 někde na stanici, tak si pro to můžeme rovnou sáhnout a pracovat s tím. A pak máme druhý přepisovač, který už by měl používat umělou inteligenci, takže je tím pádem o něco chytřejší, výsledky z toho jsou kvalitnější a ten právě zavedl Honza Cibulka. Udělal nám do něj přístup a ten je mnohem lepší a využíváme ho podle mě všichni a denně i desetkrát. Ten využíváme hodně často.

Kdyby se ty nástroje používaly, kdo by měl být za ten obsah odpovědný? Jestli je to ten editor, nebo redaktor...

Rozhodně je to vždy redaktor a editor. Ty výstupy, ani když máme vlastně ten přepisovač už integrovaný do našeho pracovního systému, tak rozhodně... myslím, že se někdy debatovalo, že bychom tam házeli čisté přepisy, ale vždycky potřebují lidskou editaci. Ty výstupy jsou vlastně hrozně nekvalitní, je na nich poznat, že jsou strojové. Přepisují slovo od slova, zatímco editor nebo redaktor, který přepisuje rozhovory, tak vynechává všechny

ty spojovací slova, vsuvky, nadužívání různých zájmen... tohle je potřeba udělat ručně, člověkem, aby to bylo ke čtení. Takže ta věc nám sice neuvěřitelně pomáhá, zrychluje nám práci, ale vždy je tam potřeba lidské ruky. Určitě nám na desetiminutovém článku tyhle přepisy ušetří tak 20-30 minut času. Ale stejně vyprávět i krátký rozhovor třeba desetiminutový, může zabrat přes hodinu práce právě proto, že je tam hodně lidské práce při čištění textu. Třeba z 25 minut rozhovoru jsme schopni udělat článek za hodinu dvacet. Bez přepisovače by to bylo plus dalších 20 minut. A ještě je skvělá kontrola v tom, že když vidím strojově psané informace, tak vím, že jsem se nespletla, nepřeslechla. Je těžší změnit význam toho, co člověk říká, když to vidím přepsané slovo od slova a můžu to kontrolovat zvukem, než když poslouchám zvuk a sama si z něj píšu výpisky a přepisuji ten rozhovor. Je to pro mě výrazně užitečnější a lepší, když to srovnám s tím, jak se pracovalo před tady těmito nástroji. Je to vidět i v tom, že Český rozhlas má jednu jako mimořádně nespolehlivou kolegyni, která byla hodně jako hodně schopná přepisovat význam tvrzení a často psala věci tak, jak si myslela, že by měly být a interpretovala už to, co řekl nějaký host nějakého pořadu. A od té doby, co využívá ty přepisovače, tak je jako výrazně spolehlivější. Sice pořád je to velmi náročné s ní spolupracovat, ale výrazně se zlepšila i tahle problematická osoba, takže v tomhle vidím veliký význam těch přepisovačů

Když jste se učila s tím pracovat, jak jste na to šla? Měli jste nějaké školení nebo to bylo způsobem pokus-omyl...

Spíš to bylo takové jako „tady máš nástroj, hoď to do něj a podívej se, co z toho vzniklo“. Já jsem to začala používat ještě na jiné pozici v oddělení nových médií, když jsem dělala webařku pro stanici Plus. V té době, před dvěma lety se to zavádělo a doteď zaznívají názory některých kolegů, že nejenže tím člověk zleniví, ale že ty výsledky pak jsou trochu odlidštěné a že to není takové, jako když to přepisuje člověk z toho audia, kdy si vlastně musí dělat velmi podrobné poznámky a přepisovat celé citace. Je pravda, že když přijde nováček, tak máme kolegy, kteří po nich chtějí, aby se nejdřív naučili přepisovat rozhovory sami ručně a teprve potom jim nabídnou přepisovač. To proto, aby se naučili čistit ten rozhovor od různých doplňkových slov. Typicky je to třeba shoda podnětu s přísudkem, kterou ten přepisovač neumí udělat dobře nebo nadužívání ukazovacích zájmen jako tady, to, támhleto.... A když to člověk ručně neopraví, tak to tam zůstane a nedá se to číst.

A od toho je editor, který nováčky zaškolí?

Je to editor, případně kolegové navzájem, ale je to vždycky od někoho služebně staršího pro služebně mladšího.

Měl by čtenář vědět, že konzumuje obsah, který byl vytvořen pomocí těchto nástrojů? Nebo to ani není potřeba, že třeba ten zásah je tak malý.

V případě, že bychom používali obrázky od Midjourney, tak tam by to rozhodně mělo být přiznáno. Ale pokud jde o přepisovače, které jsou jenom naším nástrojem, něco jako tužka nebo klávesnice, tak tam si myslím, že to potřeba není. Že ten lidský zásah je v tom tak silný, že je to opravdu jenom výpomoc. Ale pokud by ten obsah byl z větší části tvořený umělou inteligencí, tak určitě by to podle mě mělo být přiznané. Teď je kampaň na stanici Jazz a D dur, pro kterou jsou k vidění v metru plakáty vytvořené právě ve spolupráci s Midjourney. A tam je to přímo přiznané. Je to tam tak dobře napsané, že si toho nešlo nevšimnout. A to si myslím, že je správně. My jsme vlastně o umělé inteligenci teď nedávno mluvili s kolegy právě, co se týče těch ilustračních obrázků, že poté, co nám byla smazána část databáze obrázků, jak moc vlastně s tím pracovat, když dobře zvolený ilustrační obrázek dokáže

navodit atmosféru článku. Stejně tak, jako to dokáže udělat ilustrační hudba k nějaké reportáži. Tudiž by mělo být zvýrazněné, že tohle je obrázek dělaný přesně na míru tomuto článku. Právě dneska jsem hledala ilustrační fotku s obrázkem lékaře a všechny fotky v databázi jsou univerzální v tom, že se dají využít jak ke smutné, tak i k dobré zprávě. Ať už k článku, že někde ve vesnici otevřeli novou lékařskou ordinaci a je to dobrá zpráva pro místní, tak ji může použít stejně k reportáži, že mladá lékařka si stěžuje na porušování zákoníku práce v nemocnicích. A je to ta stejná fotka, zatímco kdybych si měla vygenerovat fotku od Midjourney „lidé mají radost z nové ordinace“, tak by ta první fotka pravděpodobně vypadala jinak a měla by jiný náboj než v případě lékařky, která si stěžuje na přescasy. A v tom by to mohlo být zavádějící nebo pozměňující tu interpretaci.

Je to pro vás teda nějaký krok navíc? Obrázky generovat místo procházení nějaké databáze, fotobanky.

Zatím to rozhlas nepoužívá z nějakých jako právních důvodů. Spíš se o tom v tuhle chvíli ještě debatuje, než že by to byl krok navíc. Samozřejmě že dohledávat ilustrační obrázky je neuvěřitelně pracné a otravné, ale výhoda ilustračního obrázku z databanky je právě v tom, že se dá použít ve vícero kontextech, protože je hodně emočně sterilní. Zatímco kdybych dala do Midjourney perex článku a řekla „vytvoř mi k této zprávě obrázek“, tak že by mi do toho už vnesl nějaký náboj.

Tak můžeme jít na rekomendační algoritmy. Víte, co to jsou rekomendační algoritmy a jak fungují?

Úplně ne. To myslíte ty, které napovídají třeba klíčová slova, nebo články? Je tam nějaká sekce Čtenáři doporučují, ale já to vlastně nepoužívám. Nepotřebuji algoritmus, aby mi doporučil dobrou práci svých kolegů, protože tam jsem, když ta práce vzniká, takže cíleně si jdu přečíst to, co chci a už uživatelsky nepotřebuju, aby mi to něco napovídalo. Ale vím, že se mi tam zobrazují články, u kterých si často řeknu, že jsem je četla třeba v jiném prohlížeči, když si je otevřu třeba na mobilu. Většinou se mi tam objevují autorské články, větší reportáže... nezobrazují se mi tam třeba agenturní zprávy, ale hodně kontextové věci. Ani nevím, jak to reálně používá někdo jiný, nebo co se s tím děje dál.

Vzpomínáte si na řádek Robot doporučuje? Jaký na to máte názor, jestli by se to mělo vrátit, jestli to fungovalo, nefungovalo, jestli jste to vůbec zaznamenala.

Já jsem to zaznamenala, vím, že to tam bylo a byl u toho zábavný obrázek, ale nevyužívala jsem to. Tím, že nejsem vlastně uživatel iRozhlasu, tak nevím, jak to vnímal uživatel iRozhlasu.

Pracujete nějakým způsobem s automatickým doporučováním v rámci vaší práce, nebo vůbec...

Spíš ne. Na základě nějakých štítků, popisů, nám to vyhodí všechny články na dané téma, a s tím hodně pracujeme denně, protože dohledáváme články k tématu. Všechny ty články mají nějaký související článek uvnitř, takže pracujeme sice s našeptávačem, ale ne ve smyslu nějakého složitějšího našeptávání, ale je to jenom katalogizace. To používáme denně asi tak čtyřikrát.

V jakých situacích to používáte?

Bud' se chci podívat, co jsme o tématu psali anebo chci doplnit ten svůj text textem na podobné téma. Ale právě ve smyslu katalogizace. To je jenom chronologické seřazení od

nejnovějšího k nejstaršímu článku na dané téma. My tomu říkáme asety a jsou to ty vložené články, které se zobrazují v jiném článku, anebo pak na konci. Ale to bych ani neřekla, že je úplně umělá inteligence. A myslím si, že i za tím výběrem je dost lidské práce. Pořád je za tím lidský faktor, který to kontroluje. Třeba my tam máme navrhovač štítků, který by mohl tyto věci katalogizovat sám, ale často ho nevyužíváme, protože neumí moc pracovat s nadřazenými slovy. Pokud budeme mít zprávu ze zoo o pandě červené, tak úplně nedokáže říct, jestli je to ohrožený savec anebo zvíře na seznamu ohrožených zvířat, ale spíš mi to napíše ty konkrétní slova, která se v tom článku objevují, což jsou prostě zoologická zahrada, panda červená, Liberec... Už by tam chybělo právě to, co potřebuji, jako ohrožené zvíře, export zvířat, obchodování se zvířaty a tyhle ty související a nadřazené pojmy.

Dokážete si něco v tomhle směru, v této oblasti představit, že by vám tu práci ulehčilo? Protože to zní docela dost časově náročně.

Je to otravné a časově dost náročné. Kdyby ty nástroje byly chytřejší a víc toho suplovaly, přizpůsobily se, tak ano, ale v tuhle chvíli ty nástroje, které máme k dispozici, takové nejsou. Pořád je člověk efektivnější.

...takže byste potřebovali něco, co by líp chápalo kontext těch článků a líp by psalo ty tagy, jestli tomu dobře rozumím...

Přesně tak.

Kdybyste si představila, že při editování by vám algoritmus nabídl dalších 10 článků a vy byste z toho vybrala ty, které by se doporučily čtenáři podle jeho zájmu... nebo to dělá ten našeptávač?

No, dělá i nedělá. Dovedu si představit, že by si třeba ten nástroj dokázal sám říct, že teď píšu o psychickém zdraví mladých lidí a rovnou by mi to nabídlo související články o psychickém zdraví mladých lidí bez ohledu na to, jaké konkrétní štítky bychom k tomu přiřadili. Protože teď je to tak, že se musím snažit trefit do představy o tom, jaký štítek použil kolega přede mnou a některé tagy se nám třeba opakují, nebo jsou obdobné. Například máme téma změna klimatu a klimatická krize a bylo by skvělé, kdyby umělá inteligence dokázala pochopit, že když píšeme o tom, že je rekordní teplo v říjnu, tak teď je to tak, že zadám štítek rekordní teploty, ale pak musím zadat štítek klimatická změna, aby mi to našlo další kontext a potom musím zadat měření českého hydrometeorologického ústavu a vybrat další věci. Kdyby ta inteligence dokázala všechny tyto věci spojit do jediného kontextu, tak by to bylo samozřejmě hrozně užitečný a rychlý. Ale zase by to musel vždycky vybrat člověk a editorsky rozhodnout. Vždy je tam na konci ten člověk. Ten editor, který rozhodne, že chce, aby článek o tom, že je rekordně teplý říjen, byl rámován klimatickou krizí. Protože to může být klidně rámováno jenom tím, že je hezké počasí. A to je vlastně ten lidský přínos, který by tam neměl zmizet. Vlastně se v tom editorsky lišíme. Jsou kolegové, které by ani nenapadlo k článku o tom, že v říjnu je 17 stupňů, k tomu třeba přidat štítek klimatická změna. A já bych to tam určitě chtěla dát.

Takže je to hodně o tom, že vy teď musíte mít velký přehled kdo co napsal, jaké jsou ty články...

Je to tak, že jako editor radím směnářům, kteří chystají ty věci, nezapomeň tam dát tohle téma, protože o tom psal ten a ten minulý týden a já jsem u toho zrovna byla, když jsme to řešili a byl to dobrý článek a teď se tam perfektně hodí. Je to hodně o lidském přístupu, o editorovi.

Ted' zase taková teoretická otázka, ale když bychom si představili, že by vznikla nová stránka iRozhlasu, která by byla čtenáři ušitá na míru, podle toho třeba co on konzumuje, že by vedle toho hlavního webu existoval nějaký takový personalizovaný, co byste na to řekla?

Určitě by to bylo vítané, protože mě konkrétně obtěžuje, když si musím prolistovat sportovní články, abych se dostala k článku o vědě, protože prostě tahle jedna rubrika je pod druhou rubrikou, nebo jsou dny, kdy sama mám chuť číst víc věci o polských parlamentních volbách, ale jsou dny, kdy si otevřu web a nechci tam vidět třeba ukrajinskou válku, ale raději bych si četla něco z lifestylu. Takže bylo by skvělé si to nastavit. A vlastně ta stránka je tak rozřešena, že mi přijde, že je i snadné si jí personalizovat už teď, když vím, v jakém pořadí je tam nejvíce relevantních 9 hlavních textů, a dalo nejvíc práce je napsat, ale potom už jsou ty rubriky. Takže já se můžu selektivně proscrollovat rubrikou ukrajinská válka a útok Hamásu, abych se dostala například k fotbalu. Takže myslím si, že nějak už to jako funguje, ale kdyby to bylo extra, tak vlastně proč ne? Ale samozřejmě je tady pak nějaká etická otázka jako prohlubování sociálních rozdílů, bublin a toho, že vlastně vidíme svět selektivně. Přála bych si, aby tam proběhla nějaká hlubší debata o tom, jestli chceme být účastni tohoto světa bublin.

A jak to vlastně funguje teď? Kdo určuje, jaký ten článek a kde bude?

Ted' je to editor, který řeší těch hlavních devět článků. Potom je to zase editor, který řeší umístění rubrik podle nadřazenosti, to znamená, že jsme schopní nadřadit rubriku Útok Hamásu nad rubriku Ukrajinská válka. A je to zase editor, který pak tam dá nějaké výběrové rubriky, jako Můžete si přečíst nebo Doporučujeme na základě toho, který článek mu přijde nejvíce relevantní. Ale potom jsou rubriky, které se řadí chronologicky, ale náš systém je vlastně hodně personalizovaný, a i v rámci rubrik bych mohla něco stáhnout a něco víc vytáhnout.

Jako příklad, před 2 týdny jsme měli rubriku Útok Hamásu a chtěli jsme na čtyři dny udržet článek nahoře přehledně, kdo je kdo na blízkém východě a teprve pod ním se řadily ty další zprávy od nejnovější po nejstarší.

Kdyby řekněme iRozhlas fungoval tak, jak jste už zmínila, kdyby to bylo personalizované, jaké výhody a etické výzvy by to přineslo tomu čtenáři a Vám jako editore. Jestli by to nějak změnilo Vaši práci.

Čtenář by rozhodně měl rychlejší přístup k tomu, co ho zajímá. A na druhou stranu, někdy mě samotnou zajímají věci, o kterých bych dřív nevěděla, že by mě mohly zajímat. Často říkáme, že o čtenosti článku rozhoduje titulek, ale pokud by to bylo personalizované, lidi by si rozhodovali podle témat, tak ani dobrý titulek by nespádl do jejich selekce. A přesto zprávu ze sněmovny můžete napsat strašně nudně ale i hrozně zábavně, ale já si nastavím, že nechci vůbec číst zprávy ze sněmovny, byl by tady určitý problém. A editorsky by to vedlo k tomu, že bychom dělali míň zábavné titulky s tím vědomím, že kdo si nenastavil zájem o sněmovnu, tak si o tom nepřečte nic, ani kdyby se tam Okamura svlíkl do trenek. Koho by tedy nezajímala domácí politika, tak by se tuto zprávu nikdy nedozvěděl. Dovedu si představit, že by to bylo samozřejmě hrozně populární, protože vybrat si to, co si chci číst a vyhnout se tomu, co číst nechci by bylo hrozně pěkný a lákavý. Mě samotnou by to asi lákalo. Protože nikdy nechodím na sociální sítě, protože nemám náladu prodírat se různými druhy obsahu, když chci vidět konkrétně jen jeden. Dejme tomu je neděle ráno a já nechci vidět, co udělal Hamás, ale zajímá mě, co dělají různé popové celebrity. Ta nabídka se tomu

sice hned přizpůsobí, ale pak už to není úplně užitečný nástroj, když o hodinu později potřebuji vidět ten Hamás... ale jsou chvíle, kdy nechci vidět nic. Pokud by to šlo nastavit jako forever, tak by to byl velký problém.

Zajímá mě, jestli takhle může fungovat zpravodajský web veřejnoprávního média...

V tuhle chvíli asi nemůže. Já si myslím, že ta vnitřní debata v rozhlasu je nastavená tak, že máme informovat o všem a nabízet všeobecný přehled. Takhle daleko v té debatě ještě nejsme, abychom vůbec připustili to, že můžu mít vedlejší záložku, vedle hlavního okna, kde si můžu naházet, co mě baví, aby mi to neuteklo je něco jiného, než kdyby to bylo hlavní okno. To je hodně jiné. Ve vedlejších okně je to v pohodě. Už teď, myslím, že v mobilní aplikaci tam tu možnost máme a já jí třeba využívám na mujRozhlase. Když jsem například večer doma a nechci poslouchat politické rozhovory, ale četbu, tak tam mám naházené různé pořady přímo nabízející četbu na pokračování a když si otevřu mujRozhlas, tak jdu rovnou do té četby a nemusím se dívat, co je nového na úrovni reportáží a politických debat. Tam to určitě má smysl, ale musela jsem si to sama vytvořit. Musela jsem si to sama zaklíknout, že chci.

Pokud se vrátím k doporučovací algoritmu, měl by čtenář vědět o tom, že vlastně konzumuje obsah, který je mu doporučován ne přímo člověkem, ale nějakým algoritmem a pokud ano, jak by to mělo být označené?

Určitě. Ten rozdíl mezi „tohle se nám opravdu povedlo, zkuste se na to podívat, i když byste si to na první dobrou nevybrali“ a mezi „tohle je na podobné téma, mohlo by vás zajímat“, je veliký. To jsou vlastně dvě úplně mimoběžné věci a rozhodně by tam mělo být zdůrazněné, že jedno jsme vybrali pro vás my, a to druhé pro vás vybral stroj, protože se vám to zpravidla líbí. No a označené by to bylo někde v názvu té rubriky nebo zdůrazněné hned pod ní. Tohle se vám zpravidla líbí, a tak vám to automaticky doporučujeme a tohle doporučujeme my sami, protože jsme na tom vytvořili velký kus práce. Protože buď je to něco, co vás udrží v jakékoliv bublině, kterou si vytvoříte, anebo to je něco, co víme, že to mělo ohlas, stálo to extra moc peněz a bylo to extrémně náročné vytvořit nebo je to v něčem unikátní. A přestože to pochází z jiné zájmové kategorie, z jiné bubliny, tak pořád si myslíme, že to za to stojí. Jsou to dvě rozdílné kategorie a já to třeba hodně vidím na Spotify. Například když si pustím nějakou kapelu, tak mi to nabízí další podobné kapely a já se vnořím do podobného vibu, zato když si otevřu recenzi třeba a vidím tam co doporučují, tak mě to zase dostane do jiných vod, což je osvěžující.

Ted' už mám asi poslední otázku, jestli byste měla zájem o nějaké školení nebo manuál ze strany projektových manažerů, kteří zvažují zavádění inovací.

Určitě, určitě by o tom měla proběhnout debata. Měli bychom to zkusit vymýšlet společně. Protože rozhlas je veliký a dělá spoustu věcí a nikdo neví dost dobře, co dělá jiné oddělení.

Co by vám v tomhle směru nejvíc pomohlo?

Právě debaty s techniky. Aby ten, kdo rozumí technické stránce věci, mluvil s tím, kdo je víc editoriální a aby se ty světy propojovaly. Historicky vzato, než ještě vznikl iRozhlas jako samostatná značka, tak jsme tady měli redakční systém, který byl neuvěřitelně nepřátelský vůči psaní textu. Pro každý odstavec měl samostatné okno a mezi nimi se muselo proklikávat. Takže jste nepracovali v jednom textu, ale v několika samostatných okýncích, což bylo neuvěřitelně nepraktické pro fungování textu, jak má ten text plynout. Pro každý enter, jen aby se mi to lépe četlo, jsem musela vytvořit nové okýnko. A stejně se to používalo,

protože někomu z technického oddělení přišlo jako dobrý nápad, že by se takhle mohly psát články. Pomohlo by nám, kdyby nám technici naslouchali, jak my s tím pracujeme a nechali si poradit, co v tom potřebujeme. Prostě nepotřebujeme mít okno pro každý odstavec, ale potřebujeme něco, co nám bude ulehčovat práci. A ukázat jim, jak ta práce probíhá krok za krokem.

V souvislosti s tím, co říkáte, napadá Vás už teď konkrétně, které kroky by šly zlepšit? Jsou to třeba ty tagy, o kterých jsme se bavili?

Ano, například jako v případě přepisovače, který zavedl Honza Cibulka, který je sice redaktor, ale zároveň je velmi technicky zdatný, takže věděl, že máme přepisovač, byť ho sám nepoužíval, protože když píše články, tak pracuje jinak, než aby přepisoval velké kvantum textu. Ale slyšel, že tohle funguje a že mi to používáme pro zlepšení práce, tak zavedl něco, co funguje ještě trochu líp. Takže to je třeba velká fúze tímto směrem. A pokud budeme mluvit o obrázcích, tak bych chtěla nějakou velkou etickou debatu. O tom, kam vás může vytváření ilustračních obrázků vést, co všechno vlastně obrázky můžou změnit? Že to je hodně podobné, jako když televize Nova začala dávat do reportáží náladovou hudbu. Stejně tak dokáže i obrázek umocnit charakter zprávy. Třeba jako když našemu fotografovi strašně záleží na tom, aby když píšeme o počasí, tak aby na fotkách byli lidé oblečení úplně přesně podle počasí, které má být dnes. Takže i tyto věci jsou důležité.

A v rámci toho doporučení, Vás něco napadá?

Tam bych si taky přála nějakou etickou debatu. Máme sice skvělý nástroj, který každému čtenáři hrozně ulehčí vyhledávání jeho oblíbených témat, ale otázkou je, jestli ta představa editorů o tom, co by čtenář měl číst, nejde do pozadí oproti demokratizaci toho, že si chci vybrat svůj vlastní obsah. Hodně často mluvíme o tom, co udělají sociální sítě se společností. Jak se každý může uzavřít do své bubliny a zůstat tam neomezeně dlouho. Jestli se nestáváme spíš sociální sítí než zpravodajstvím. Určitě bych v tomto směru uvítala nějakou debatu na úrovni, ale aby to nebylo pouze jednorázové dvouhodinové školení. Chtěla bych, aby to byla dlouhodobá debata, proces.

Respondent Martin Prachař, datum 24.10. 2023

Víte, co jsou to rekomenční algoritmy a jak fungují?

Jak fungují nevím, vím asi, co jsou doporučovací algoritmy.

A jak byste to popsal svými slovy?

Jako věc, která doporučuje články k přečtení čtenářům.

Vzpomínáte si na řádek „Robot doporučuje na iRozhlasu“?

Ano.

Tak co si o něm myslíte?

Že to je dobrá věc, ve smyslu, že tomu člověku, který scrolluje tou page, tak mu to nabídne něco, co by ho mohlo zaujmout. Myslím si, že časem bude ještě víc a víc takových věcí.

Pracujete teď s automatickým doporučováním teď při vaší práci nebo ne?

Já osobně ne. Myslím si, že tam teď nic takového neprobíhá.

Kdybychom si teoreticky představili, že by vznikla stránka iRozhlasu jako třeba záložka, která by čtenáři byla ušitá na míru, tak co by to podle vás přineslo za výhody nebo etické výzvy?

Ježiš, to já nevím. Výhoda by byla, že by to tomu člověku na základě nějakých jeho preferencí nabídlo to, co ho asi zajímá, nebo by si to sám asi sestavil nějak. A etickou výzvu? Nevím, to si neumím úplně představit, proč zrovna etickou výzvu.

A uvítal byste to by jako uživatel nebo jako editor?

Jako editor asi úplně ne a jako uživatel asi taky moc ne. A to asi z toho důvodu, že bych nechtěl, aby mi něco někdo doporučoval. Že pak má člověk dojem, že s vámi někdo hraje nějakou hru a že to není úplně jako otevřené. Člověk si radši tak brouzdá a hledá si ty věci sám, než aby vám je někdo servíroval na základě něčeho.

V komerčních médiích se díky těm rekomenčními algoritmům snaží třeba zvýšit personalizaci, tak jestli by podle vás mělo používání těch doporučovacích algoritmů, opodstatnění v Českém rozhlasu jako v médiu veřejné služby? Proč ano, proč ne?

Tohle je trochu složitější, protože v komerčních médiích to dělají proto, aby zvýšili počet kliků, na kterých jsou závislí, aby mohli prodat více reklamy. U nás to tady podle mě takhle nemůže fungovat, protože některá zpráva, která je třeba důležitá, tak prostě není z povahy věci někdy třeba tak čtená nebo zajímavá, ale je důležitá. A když by se to řídilo jenom těmadletěma algoritmama, tak oni jí nacpou nahoru, dopředu, ale nebylo by to asi dobré, takže si myslím, že se tohle zrovna tady v rozhlasu asi úplně nechytne. Nemyslím si, že by se tohle tady prosadilo.

Měl by čtenář vědět, že konzumuje obsah, který mu doporučují algoritmy? Jak byste mu to případně oznámil?

No asi ideálně „Robot doporučuje“. To je docela dobré oznámení. Protože si myslím, že se to jako nestane, když se podívá člověk na iRozhlas, tak ta home page je sestavovaná lidmi a asi tak ještě dlouhou dobu zůstane. Myslím si, že tohle nebudou dělat roboti. Takhle na to doporučení je to, si myslím, dobré a název „Robot doporučuje“ je docela fajn.

Ještě k té editorské práci, když byste si představil, že když něco editujete nebo dáváte

na web nějaký článek, tak že by vám k tomu bylo automaticky nabídnuto třeba deset podobných článků, tak pomohlo by vám to nějak?

Mně jako editorovi? Jestli myslíte do takových těch souvisejících článků, jasně, to by bylo fajn. To je zrovna věc, u které by se tyhle věci jako hodily a myslím si, že to probíhá a že to bude probíhat čím dál tím víc. Takovéto automatické doplňování nebo překlápění věcí z vysílání na web, to jsou věci, kdy to bude umělá inteligence dělat celé vlastně. Ale vždycky to bude muset někdo zkontrolovat. I když zrovna v tomhle to člověk dělá, jenomže to dělá s mezikrokem. Musíte si to někde vyhledat. Když editujete článek, chcete tam dát nějaké podobně tematické věci, tak si je musíte vyhledat. Když by vám to automaticky něco jako nabídlo, tak fajn. To je usnadnění práce z mého pohledu.

Jak vypadá ta vaše práce teď a jak by jí to mohlo v budoucnu změnit?

Mojí práci to podle mě nezmění v dohledné době. Nebo si to nedovedu představit, protože vždycky musíte ten text kontrolovat. Když vzniká článek, tak já se podílím na tom, když vzniká. Jaké bude téma, jak se zpracuje... a to je všechno věc, kdy tam potřebujete člověka. Pak to musíte zkontrolovat, přečíst, vyhodnotit, a to bude zase furt na lidech. To, kde se to může prosadit, je v té automatizaci takových věcí, které dělají teď lidi, ale zabere jim to půl hodinu, ale je to vlastně jako neustále se opakující věc. Tak to bych řekl, že můžou dělat ty umělé inteligence. Tam to pomůže. Ale to už jakoby probíhá v tom rozhlase. A je to něco jiného než ty doporučovací věci.

V čem to je něco jiného?

Doporučení je něco, co funguje třeba na sítích, že to člověku neustále nahazuje nějaké věci, o kterých si myslí, že vás to bude bavit a tohle je spíš taková zajímavůstka. Já si myslím, že to je něco jiného, než když umělá inteligence nebo tyhle věci překlápí třeba příspěvek z vysílání. My ho dneska musíme přepsat, vybavit fotkou, takovou tu šablonku tam dát... dát tam domicil, dát tam autora a všechny tyhle náležitosti, související články a člověku to zabere půl hodiny nebo hodinu, podle toho, jak je to dlouhé. Tohle všechno může udělat umělá inteligence a udělat to za pět minut třeba. Pak to někdo bude muset zkontrolovat. Ale to je taková činnost tvůrčí jenom trošičku.

Je to spíš rutinní věc?

Ano, rutinní věc. Přepis vysílání do článku, tak to je věc, kde to ohromně pomůže. Ono už se to tak posouvá vždycky o krůček. Ale zatím ještě nejsme tam, že by to zvládlo celé samo.

Měl byste třeba zájem o nějaké školení ohledně těch doporučovacích algoritmů třeba od projektových manažerů a manažerek?

Proč ne a tohle zrovna probíhá. Tím, že zrovna s naší projektovou manažerkou sedíme v jedné místnosti, tak jsme furt v kontaktu v tomhle. Takže tam si myslím, že jsme relativně informovaní o tom, co se děje i co se týče doporučovacích algoritmů.

Kdyby iRozhlas fungoval jako Netflix nebo Spotify, když to hodně zjednoduším, myslíte, že by vám to něco přineslo ten princip doporučování? Jestli zpravodajský web takhle může fungovat?

Já si myslím, že ne, že to takhle ve zpravodajství úplně nefunguje. Tam jsou prostě věci, které se dějí v tu danou chvíli a není to jako... To možná v nějakém lifestylovém magazínu, kde vám to bude doporučovat zpětně věci, u kterých je ale jedno, jestli vyjdou na jaře nebo na podzim. Ve zpravodajství si to nedokážu úplně představit nebo nevím. Ten čtenář asi

nemůže mít pocit, že mu jako někdo servíruje jenom jako nějaký výsek věcí. Myslím si, že od iRozhlasu čekají všeobecnost a tu aktuálnost. Tam ty doporučovací algoritmy asi nejsou potřeba v tomhle.

A kdybychom se bavili o tom, že by vznikla nějaká jako další stránka, řekněme jako záložka, která by takhle fungovala? Nemyslím, že by celý iRozhlas byl postavený jenom tady na tom.

Jo, že by si to na tom jako někdo přehodil. Nevím. Mně osobně by to asi nevyhovovalo a myslím, že tím směrem se ani nevydáváme.

Respondentka Nela Vejvodová, datum 24.10. 2023

Začneme rekomenačními algoritmy. Víte, co to je a jak fungují?

Myslím si, že vím, co to je. Podle mě algoritmy vyhodnocují to, na co se na internetu díváte, na co klikáte a podle toho vám doporučí obsah, který si myslí, že je nejbliž tomu, co už na internetu děláte, a to kvůli zvýšení dosahu, abyste tam strávili co nejvíc času.

Vzpomínáte si na řádek „Robot doporučuje“ na iRozhlasu?

Ne.

A ani jste to nezaznamenala jako uživatelsky?

Ne.

Pracujete teď nějak s takovýmto doporučováním v rámci vaší editorského práce, nebo ne?

Máme tam sekci *Doporučujeme*, která se ale neřídí nějakou AI nebo robotem, alespoň o tom nevím. Spíš ten základ je takový, že předvýběr článků, které se pak v *Doporučujeme* objeví, dělá editor klasicky. To znamená, že jde o nějakou naši původní zprávu, nebo o něco, co se během dne fakt hodně četlo a my to chceme do dalších dnů ještě víc znova napumpovat, takže to dáváme do takové speciální tabulky, ze které se to pak sosá. Předpokládám, že tam nejsou jiné články než ty, které mi tam vybereme. Takže takhle s tím pracujeme.

Jak vypadá ten váš výběr doporučených článků konkrétně, ta vaše práce?

Je to na základě toho, že když jde o zprávu, na které se strávilo fakt spoustu času, chceme ji ještě znovu prodat i když už nemá místo na homepage, což znamená druhý den, nebo ono je to občas i po půl dni. Typicky to může být třeba „Čapí hnízdo“, které se někdy ne úplně čte, protože to je tak složitá kauza, že se v ní už lidi moc neorientují. Ale na druhou stranu Terka Čemusová na tom dělá opravdu dlouho a už podle mě zná každé písmenko v tom spisu, tak je prostě škoda ještě znovu neotočit ten článek. Dám příklad, kdy článek od pěti ráno držíme na homepage třeba do čtyř odpoledne a v ten samý moment to dáme jednak do rubriky *Domáci*, ale zároveň to zařadíme do rubriky *Doporučujeme*, takže je to vlastně vidět pořád na homepage nahoře, kde je takzvaný *b sloupec*. A ještě někde, podle mě zhruba ve druhé čtvrtině naší stránky, je zase řádek *Doporučujeme*, kde jsou právě tyto původní zprávy, na kterých se strávilo mnoho času, anebo věci, které se hodně čtou. Není to třeba *četka*, ale jsou to věci například z vysílání, rozhovory... Včera se strašně četla ta „Bečva“, kde hostem byl hydroekolog z Plusu, tak to jsou přesně ty věci, které my tam chceme prodávat co nejdéle, takže proto je tam dáváme a takto s tím pracujeme.

Váš výběr tedy stojí na vaší znalosti všech těch obsahů?

Přesně tak a zároveň víme, jak se které články čtou, protože tam vidíme v reálném čase čtenost. Na to je i takový speciální nástroj Chartbeat. Takže to je to kritérium, které rozhoduje, ale výběr je nakonec opravdu na nás. Můžeme si tam dát cokoliv, jen si pak musíme umět obhájit, proč jsme to tam dali.

Jaká jsou ta kritéria, podle kterých se osobně rozhodujete?

Je pravda, že občas, když mám třeba dva články, mezi kterými se rozhoduji, tak tam většinou dávám zahraničí, které mě hodně zajímá, anebo různé technologické články, protože se snažím, a to hlavně o víkend, tu nabídku trochu naředit v tom smyslu, aby tam nebyly

například jen vážné politické věci. Dala bych tam něco, co je zajímavé, ale splňuje to pořád to kritérium čtenosti. Já osobně se snažím dávat do rubriky *Doporučujeme* mix různých žánrů, témat, rubrik, aby i čtenář, který to vidí jednou za čas věděl, že máme široké spektrum článků na různá témata, které se jim snažíme podat.

Když bychom si teď teoreticky představili, že by vznikla nová stránka iRozhlasu, třeba jako záložka, která by byla ušitá na míru čtenářům, a kde by to doporučování probíhalo automaticky, tak co byste na to řekla?

Z uživatelského hlediska, asi proč ne... Pokud je člověk zvyklý na tenhle ten styl konzumace zpravodajství, že mu to v jeden moment, jako na Facebooku, začne už potom předhazovat jenom zprávy z té jeho bubliny, což většina lidí to takhle má a vědomě z toho nevystupuje, tak si myslím, že z tohoto hlediska to asi bude funkční, užitečný, oblíbený. Trošku bych se bála toho, že to nebude využívaný, pokud by to byla jenom záložka, nebo kdyby to nebylo na první dobrou někde vidět, protože každý klik navíc prostě znamená, že přijdete o nějaké procento lidí, který už ten krok prostě neudělají. Takže když si to představím z toho produkčního hlediska, tak za tím vidím řádu práce, která by se myslím nevyplatila, ale z hlediska editorského asi vlastně proč ne... Myslím si, že to by byl docela zajímavý inovativní prvek toho webu, ale možná by bylo dobré kontrolovat, co přesně se tam ukazuje. Jestli to už potom nezachází do té velmi úzké bubliny a nevynechává to právě něco, co by člověka úplně nemělo minout. Protože na iRozhlas jde s nějakým důvodem, že si prostě chce přečíst, co se děje a není to ten typ webu, kde dostáváte opravdu jenom úzkou škálu témat a názorů.

Proměnilo by to nějak vaši práci? Mluvila jste o tom výběru... Tak jestli by se pro vás něco změnilo.

Možná by to bylo trošku víc práce pro nás, že bychom museli přemýšlet ještě nad tímhle aspektem, ale myslím si, že by to asi šlo paralelně s námi.

Jaké výhody nebo etické výzvy by to přineslo?

Já si myslím, že největší výzva asi by byla neuzavřít toho člověka v té velmi úzké bublině a pořád mu připomínat, že máme i jiný obsah, který taky stojí za to a že klidně, ať se kouká na svoje oblíbená témata, ale ať občas dá šanci i tomu proudu mimo.

Máte zkušenosti i ze Seznamu, kdybyste to mohla i možná srovnat, tak podle Vás má používání těchto algoritmů opodstatnění v Českém rozhlase, ve zpravodajství veřejnoprávní služby?

Myslím si, že ne. Právě proto, že na Seznamu to v určitých dnech už vedlo do takového extrému, že kolega, který seděl vedle mě, tam měl úplně jiné články. Jednak na Seznam.cz, kde jsou nejdřív Seznam zprávy, Novinky, jednak i dole, kde jsou zprávy z jiných médií, kde měl taky úplně jiné články. V jeden moment už to bylo tak rozdílné, že když jsme potřebovali jeden článek dohledat, on toho nebyl schopný, protože ho vlastně vůbec neměl ve svém výběru a z toho hlediska, že jsme ho potřebovali k práci, to bylo opravdu nepříjemné. Na druhou stranu uznávám, že speciálně novináři s médii, jako jsou články, zacházejí velmi rozdílně, než jako normální čtenáři, ale myslím si, že do veřejnoprávní služby to nemá smysl dělat. Ta by naopak měla mít celé to spektrum, a právě by nemělo docházet k tomu, že bude na výběr jenom určitý segment a bude velmi těžké se doklikat k něčemu, co není v té vaší oblíbené sortě. A taky si myslím, že vlastně Seznam to vyhnal do extrému i kvůli dobrým příjmům. Víím, že se u některých článků pak násobně zvýšila čtenost.

Teda nebylo to u Seznam zprávy nebo Novinky, ale spíš u těch lifestylu jako Pro ženy, Auto a podobně, u kterých vlastně tady to doporučování podle obsahu vytáhlo hodně tu čtenost, když už to tam lidem potom jenom přihazovalo recept na buchtu, pak už čtyři další recepty a ty lidi na to klikli. Takže to vlastně velmi pomohlo spíš lifestylovějším sekcím Seznamu a myslím si, že se to, upřímně, v některých momentech zvrtilo. Že ten Seznam upřednostnil spíš ty příjmy, které z toho má na úkor toho, že rezignoval na to, že tam zajistí aspoň nějakou elementární úroveň, pod kterou už to doporučování nepůjde, aby tam nezmizely nějaké důležité věci. Nemám to empiricky ověřené, ale jak jsem se tam několik let pohybovala a jak jsem s tím pracovala, mám dojem, že se prostě tohle stalo.

A dalo by se tomu nějak zamezit?

Určitě dalo, ale muselo by to být exekutivní rozhodnutí někoho, kdo řekne, že peníze jsou fajn, ale za nějakou hranici toho, co lidem budeme nabízet, už nepůjdeme, protože chceme mít nějakou základní úroveň článků, která tam být má.

Mělo to nějaký vliv na podobu titulků?

To doporučování ne, tam spíš titulky, pokud se měnily, tak to bylo proto, že se třeba něco nečetlo a zkoušelo se, jestli se to bude číst líp s jiným titulkem. Ale to dělali opravdu editoři a bylo to rozhodnutí člověka a nikdy to neprobíhalo s nějakým programem. To ne.

Měl by čtenář vědět, že konzumuje obsah, který mu byl doporučovaný algoritmem?

Určitě.

Jak byste mu to případně oznámila?

Určitě by to měl vědět už jenom z toho důvodu, že podle mě pořád spousta lidí neví, že to například na Facebooku takhle funguje a na Tik-Toku je to dokonce úplně extrém. A pak se to vyhrtí úplně do radikálna, protože jsou už jen v té svojí omezené bublině a nedojde jim, že existují i jiné zprávy. Takže určitě by na to mělo být nějaké upozornění. A asi bych to někde dávala. Samozřejmě, pokud bychom se bavili o tom, že bude k tomu založena speciální sekce, tak to bych tam dala velmi výrazně. Ať se to klidně jmenuje třeba *Robot doporučuje*, ale právě pod tím bych dvěma větami nebo klidně nějakým headlinem, vysvětlila, že jde o články, které vám tady nabízíme na základě vaší aktivity. Za mě by bylo dobré to mít i v každém tom jednotlivém článku, který byste si rozbalila právě na základě tohoto robotického výběru. Třeba někde dole jako poznámka ve smyslu, že „tenhle článek byl vybrán, protože je to vaše krevní skupina, ale venku jsou i jiné zprávy“.

Kdybyste si představila, že při editování by vám bylo nabídnuto třeba automaticky dalších 10 článků, pomohlo by Vám to nějak při výběru?

To by mi pomohlo hodně. My máme na iRozhlase takové dvojí související skupiny textů. Tou první jsou vložené články přímo do textu. Myslím, že se to označuje jako *Čtete také* a zároveň na konci, když si zascrolujete až dolů, tak jsou tam 3-4 články pod názvem *Související*. Rozdíl mezi tím není. Ptala jsem se a vlastně není žádné kritérium, podle kterého se rozhodne, jestli se článek objeví v *Čtete také* nebo *Související*. A tak je občas dost složité ukočírovat, co je ještě jako *Související* článek a co ne. Největší problém pak je, když tady máme juniorní směnaře, nebo prostě někoho, kdo právě začíná a nemá to ještě v ruce, tak tam občas dá úplně nesmysly. Jeden konkrétní příklad: když to byla zpráva o tom, že v nějakém lese na jihu Čech si bajkeři postavili nelegální cyklodráhu, ničily tam přírodu a plašili zvěř, tak mladá začínající kolegyně tam dala jako *Související*, že se v pražské ZOO

narodilo mlád'átko gorily. Když jsem se ptala proč, ona to ospravedlnila tím, že to rovněž bylo v sekci *Příroda*, kde byl tento článek, který ona dělala. A tyhle ty věci by se vlastně velmi jednoduše odstínily právě tím, kdybych už tam měla nějaký předvýběr článků od robota nebo od čehokoli jiného, protože teď tam vlastně každý ten článek dáváme ručně a hledáme ho ručně. A do velké míry to stojí na tom, co si každý z nás pamatuje, například že už to vyšlo, ale i tam jsou nějaká pravidla v tom smyslu, že ten článek by neměl být starší než jeden rok. Samozřejmě je to téma od tématu, ale nejen stáří článku rozhoduje, ale i klíčová slova. Každý si tam píše co chce, byť už tady byla snaha to nějak normalizovat na jedno klíčové slovo. Takže hledat nějaký článek na jedno klíčové slovo, i když víte, že tam byl, najít nemůžete. A na tomto principu my vlastně hledáme ty články, které tam dát. Já to třeba dělám i tak, že si to googlim přes *Site:iRozhlas* + klíčová slova. Takže kdyby nám odpadla tady ta část práce, tak by to pomohlo, protože někdy to upřímně zabere stejně času jako ten článek do toho systému dát a najít tam jenom ty související.

Mluvila jste o nějakých těch kritériích, tak to jsou vaše vnitřní kritéria, nebo máte někde sepsaný kodex?

Ne, ne, to jsou naše vnitřní kritéria, ale obecně ono jich moc není. Mělo by to být vždy kontextově relevantní. Například ve zprávě, která už má pokračování, jako třeba teď Izrael, Gaza, tak tam už nedat článek ze sedmého října, kdy Hamás podnikl útok. To už je vlastně zastaralá zpráva. Nedáváme tam ani komentáře jako *Související*, protože tam na první pohled z titulku není vidět, že jde o komentář. A ty titulky komentářů jsou většinou dost ostré a ne vždy úplně veřejnoprávní. Jinak jde hlavně o to, aby to bylo kontextově správné a taky zajímavé. Když máte dva články k dispozici, tak si vždy vyberete ten zajímavější.

Takže je to postavené na nějakém vkusu a odpovědnosti editora?

Přesně tak.

Měla byste zájem o školení či manuál od projektových manažerů na ty doporučovací algoritmy?

Asi proč ne? Nejspíš by záleželo, do jaké míry by to pak bylo propojené právě s tou naší denní prací. V ten moment jsem všema deseti pro, protože já tyhle věci potřebuju. Je určitě zajímavé zjistit, jak to funguje právě zevnitř a třeba i na základě toho pak upravit naši práci. To si dovedu představit.

Teď teda ta druhá část. Používáte například nástroje ChatGPT, Whisper nebo Midjourney při Vaší práci?

Jak kdy. Vyloženě při editorský práci ne, protože tam vlastně jako není asi důvod, proč. Nebo teď mě jako nenapadá, proč bych to dělala, ale když třeba jako píšu, protože občas ještě nějaký články jako píšu, když mám čas, tak tam to používám třeba jako takovej druh brainstormingu. Ne vždycky, ale třeba jako když se zaseknu nebo prostě už fakt jako nevím, jak do toho tématu, tak si právě v CahtGPT nechám třeba jako vygenerovat, nevím prostě plácnu jako 5 hlavních důvodů izraelsko palestinského konfliktu a vím, jaký to má limity samozřejmě, ale vlastně jako pomůže mi to, že pak pak jako mě to trošku posune v tom jako jak nad tím přemejšlet, že jsem třeba na něco zapoměla, takže je to spíš takový jako jenom

uřizování myšlenek, tak to mi v tom jako docela pomáhá. Nepoužívám to často, jenom jednou, dvakrát měsíčně možná tady to jako udělám, plus taky je vlastně fajn, že tam se dá najít nebo tam, že jo napíšete jednoduchý příkaz *základní fakta o Spojených státech* a nemusím to prostě na Wikipedii hledat v strašně dlouhým textu, takže to je vlastně taky docela fajn. Midjourney nepoužíváme vůbec, ale vedla se o tom debata, co se týče ilustračních fotek. Bylo to ještě teda před hackerským útokem, jsme trpěli tím, že třeba takové ty generické věci jako *peníze, lidi na ulici* a podobně, jsme měli čtyři fotky, které jsme točili furt dokola, a to už pak prostě jako není hezký. Kor když tam přesně dáváte ty související články, který má každý tu samou fotku, tak se tak jako nadhodilo, jestli by nebyla cesta generování obrázků umělou inteligencí. Nakonec ta debata asi jako nevyústila v žádný podle mě závěr, ale bylo to něco, nad čím jsme se zamýšleli, jak to uchopit v tom smyslu jak dát lidem vědět, že to je jako vygenerované umělou inteligencí. Takže jsme to řešili hlavně v tom etickém aspektu, jak případně dát lidem vědět a jestli by to vlastně vůbec bylo fér, používat ilustrační obrázky, vygenerované umělou inteligencí. Ale jako žádný závěr tam nebyl, to ne.

Zmiňovala jste ty limity, tak jaké to podle vás jsou?

Že se tomu nedá věřit. V těch základních věcech ano, ale už jsem si taky vyzkoušela, že v určitý moment, když třeba po něm chcete, aby rozvinul téma víc a víc, tak v jeden moment, už si začne vymýšlet a velký problém je v tom, že tam pořád nedává odkazy. A když tam napíšete, *dej mi na to odkaz*, tak on napíše, že ještě nemůže dávat odkazy, což si myslím, že by ten problém docela vyřešilo. Kdyby tam dával zdroj těch informací, tak by to asi dost pomohlo. V těch základních věcech a pokynech nebo v té základní konverzaci to problém není, ale když potom chcete jít do hloubky, tak tam už mám zkušenost, že začne plácát nesmysly, ale zároveň je plácá dobře, takže je strašně těžké odlišit, že to není pravda a přijde se na to třeba jenom tak, že prostě víte, že to není pravda.

Takže to používáte víc při redaktorské než při editorské práci? Chápu to dobře?

Přesně tak.

Jak jste se k tomu používání dostala? Sama jste se o to nějak zajímala, nebo to bylo formou pokus-omyl?

Já jsem tady v tom trošku jako babička v moderních technologiích, ale na ChatGPT jsem byla opravdu hodně zvědavá. Protože jsem viděla na sociálních sítích vygenerované úryvky toho, co to dokáže napsat. Hodně se řešily testy, diplomky... První impuls byl, že jsem chtěla zkusit, jak by vypadal akademický text a jestli to bude na něm poznat. Není to poznat. Takže si myslím, že to bude hodně velký problém, protože to není něco, co by ten ChatGPT kopíroval odněkud. Ten to prostě vymýšlí, takže když to proženete programem na srovnávání, tak nic nenajdete. To byl tedy ten prvotní impuls, proč jsem to vyzkoušela a pak mě vlastně příjemně překvapilo, že to má i tady tu funkci jako „zjednoduš mi něco“, nebo

„5 nápadů na...“ a že mi to začalo docela pomáhat, když jsem se na něčem zasekla. A ostatní věci jsem nezkoušela, protože tomu nerozumím.

Kdo by podle vás měl být odpovědný za takový obsah, jestli je to třeba redaktor nebo editor, když si pomůžete tímhle nástrojem.

Já si myslím, že by to měla být kolektivní odpovědnost, ale ne v tom smyslu, že vlastně se nebude vědět, kdo přesně má odpovědnost, ale že všichni by za to měly být odpovědny. Redaktor v té první fázi by měl odpovídat za to, že použil ty informace a ideálně si je ověří. V druhé fázi by za to měl být odpovědný i ten editor, který to udělá znovu, aby se vychytali všechny možné nedostatky.

A opět otázka jako v první části rozhovoru, měl by čtenář vědět, že konzumuje obsah, který byl vytvořen AI?

Naprostě určitě, už jenom proto, že je tam prostě riziko větší chybovosti, ale i proto, že je etické a fér přiznat, že tyhle texty mají určité limity tak, aby čtenář věděl, že ten text, který mu předkládáme tak, že i přes veškerou snahu by mohl mít v sobě nějaké jako chyby, což by se samozřejmě stávat nemělo a je to na nějakou velkou diskuzi, ale určitě to nezatajovat, protože to by byla jenom cesta do pekel už jenom proto, že by se třeba přišlo na to, že je něco špatně a najednou by se to začalo shazovat na nějaký přístroj. Vlastně teď mám konkrétní příklad, protože ČTK to začala docela využívat. Nevím, do jaké míry úplně, protože oni jsou trochu tajemní ohledně toho, ale nedávno vyšel článek, anglická verze toho, že prodlužujeme na hranicích kontroly a v titulku bylo, že to řekl set austrie a mysleli ministra Rakušana. Ale protože umělá inteligence to ještě nemá jako zažité v ruce, že Rakušan nemusí být jenom člověk z Rakouska, tak to dala takhle do titulku a viselo to tam třeba hodinu a půl. Paradoxně v tom textu to bylo už správně, tak si třeba říkám, jestli jim to třeba negeneruje zatím jenom titulky, což je možné. A vím, že je to něco, s čím experimentují teď média nejčastěji, nejdřív s titulky a potom s textem, ale přesně tohle se může stát. A když takových textů budete mít za hodinu deset, tak se prostě stane, že to nepodchytíte, že nebude ve vašich fyzických schopnostech to všechno zkontrolovat.

Ještě mě napadá poslední věc, když jsem mluvila s Kateřinou Cibulkou, tak ona mi ukazovala to tagování, jak to funguje. Můžete mi taky o tom něco říct?

Není to teda úplně přímo to doporučování, jak jsme se o tom bavili, jako například je na sociálních sítích, ale spíš mi z toho vychází nějaká jako interní potřeba...

Jsou to ta klíčová slova, podle kterých potom v tom systému hledáme fotky, související články a já jsem do toho pořád ještě úplně na sto procent nepronikla. My na to máme nějakou firmu, která spravuje databázi těch klíčových slov. Takže když ta firma měla výpadek, půl dne jsme nemohli nikam zadávat klíčová slova. Já vlastně moc nechápu, proč ta klíčová slova jsou vidět v tom článku. Buď nahoře nebo dole, jsou tam vyskládané vedle sebe ta klíčová slova, která tam redaktor zadá. Já si myslím, že by vidět být neměly, protože občas tam jsou nesmysly. Protože tam jsou prostě klíčová slova: důchodce, důchodci, důchodkyně, důchodový... vypadá to opravdu divně, když tam máte 8 klíčových slov, které znamenají úplně to samé. Ale zase nemůžete je vynechat, protože se vzájemně nepropojí. Takže potom se stává, že ten článek sice byl o důchodcích, ale nenajdu ho, pokud nezadáím správné klíčové slovo. Takže primárně je to určitě pro interní potřebu, protože si ani nemyslím, že to čtenář vyhledává tímto způsobem. Myslím si, že ten systém je velmi nedokonalý a že by to chtělo na tom zapracovat. Ale vzhledem k tomu, co se tu děje, je to v prioritách tak na stém místě.

Ale určitě velmi pomáhá, když to mám alespoň otagované v základu, ale přesně v takovýchto klíčových momentech, kdy nevíte, v jakém tvaru to tam dát, tak je to pak trochu otevřenější.

A to vidíte Vy teda v nějakém systému a zároveň to vidí i čtenáři. Můžeme se na to podívat?

To bude asi nejlepší.

Respondenti Táňa Lančová (TL) a Pavel Kordík (PK), datum 24.10.2023

PK: Jsme česká firma, která vznikla asi před 8 lety, kdy jsme původně ještě jako akademici ve škole přemýšleli, co s tím výzkumem uděláme, jestli se dostane do praxe a zjistili jsme, že to za nás nikdo nekomercializuje, takže to budeme muset udělat sami. Založili jsme firmu, a pak jí postupně budovali takovým bootstrapovým způsobem, že jsme vždycky část ze zisků investovali zase zpátky do firmy a do produktů, a tak se to postupně rozvíjelo. Na začátku nás bylo asi 5, a ještě měl každý jinou práci. Teď už nás bude asi 40...

TL: Možná ještě víc, kdyby se počítaly částečné úvazky.

PK: Mám z toho strach skoro, už je nás docela dost a firma pořád roste, a zejména v zahraničí. Od začátku to bylo dělané tak, že to, co vyrábíme, je globální produkt a že je určený pro zahraniční trh. Od začátku jsme to začali budovat tak, že jsme udělali na webu jako: „tady máte náš úžasný produkt, tady máme nějakou dokumentaci v angličtině a vyzkoušejte si to.“ Od začátku jsme neměli ani žádný sales team, takže jsme museli tak postupně přes internet přesvědčovat spíš ty technické lidi, že to chtějí vyzkoušet a docela se nám to dařilo. Hodně lidí si zakládalo účet a vyzkoušeli si to, líbilo se jim to, tak to začali třeba platit sami. Doteď máme spousty lidí, kteří takhle přijdou a začnou sami platit, což je fajn. Menším byznysům to nabízíme úplně zdarma. Kdo má do 100 000 doporučení měsíčně, což znamená, že to mají někde na webu vystavené, přijde tam uživatel a na každý page views, tak většinou si od nás vyžádá jednu rekomendaci, někdy i víc podle toho, jak je ta stránka složitá. Někdy tam je třeba víc boxíků nebo to zároveň řadí třeba celé bloky. Pak může být rekomandací i víc na jednu stránku. Je spousta malých firem, které to mají takhle malinkaté. Co se týče oborů a nějakých domén, kde vlastně ten náš produkt slouží, tak to je hodně pestré. My jsme od začátku chtěli udělat univerzální doporučovací stroj. Náš první zákazník byl Jobs.cz, kde jsme doporučovali lidem práci a nebo spíš lidi na práci. Potom jsme začali doporučovat video streaming, což je úplně jiná oblast... články na internetu, potom třeba reality a tak dále, takže to portfolio e-commerce, různé market place. Portfolio těch zákazníků je široké. Po světě máme vyšší stovky různých byznysů, které Recombee používají k tomu, aby zkvalitnili svůj produkt. To zkvalitnění produktu spočívá v tom, že tam jejich uživatelé líp najdou to, co hledají. Takže je hlavně pro nějaké firmy, byznysy, kde mám toho hodně co nabídnout a mám zároveň hodně uživatelů a kde potřebuji tomu správnému uživateli nabídnout ten správný item. My tomu říkáme item, protože potřebujeme vystihnout to, že někdy je to dům v San Franciscu, někdy písnička na Audiomacku, někdy je to článek na iRozhlase. Pokaždé to může být jiná věc, co se doporučuje.

TL: Já bych ještě doplnila, co říkal Pavel. Samozřejmě máme nějaký produkt, který je univerzální, ale není univerzální v tom, že bychom každému tomu zákazníkovi z těch jednotlivých, říkáme tomu třeba domény, což je například news, nabízeli stejnou sadu těch algoritmů, protože každá ta oblast je velmi specifická. Takže byť z toho mohlo vyznít, že máme jeden univerzální rekomendační systém, který si prostě nasadíte, ať už prodáváte domy nebo píšete blog posty na nějakou stránku... tak to tak není. Máme nějakou sadu univerzálních logik, které nejlépe vyhovují jednotlivým doménám. A protože ty požadavky mohou být velmi specifické, tak pro naše prémiové zákazníky jsme potom schopni ty logiky a ten způsob, jakým se doporučuje, ještě více vyladit na základě jejich očekávání.

V jakých médiích třeba působíte nebo na jaké bázi jsou ty spolupráce?

TL: My úplně nemůžeme mluvit o těch konkrétních zákaznících, u kterých nemáme podepsanou case study, že o nich mluvit můžeme, takže třeba na našich stránkách najdete

case study k Primě, kde je to už odsouhlaseno. Ale obecně můžeme říct, že v Čechách doporučujeme na hodně médiích, a i v zahraničí doporučujeme na hodně médiích, a jsou to velká média. Naše doporučování se vyplatí všude. Benefitů je samozřejmě vícero a máte různou sadu metrik, kterými ten úspěch můžete měřit a obvykle, když je to doporučování správně uděláno, tak vám stoupnou všechny ty metriky, protože jsou spolu nějakým způsobem provázané. Takže ať už je to větší proklikovost, ať už je to větší retence těch uživatelů, nějaké stickiness a tak dále, tak tohle všechno se správně udělaným doporučováním stoupá.

Může to fungovat i pro někoho jakoby malinkého, kterému poskytneme tu defaultní sadu logik, a ještě lépe to funguje tak, když je to nějaká velká společnost a věnují tomu obě strany nějakou péči. I malí zákazníci od nás mohou něco získat, protože nabízíme jednoduchý widget, který si můžou nakonfigurovat. Těch doručovacích míst tam máme několik, protože vlastně každé to doporučovací místo se může řídit svojí speciální logikou.

Co konkrétně zajímá vás?

Zajímala by mě ta konkrétní spolupráce s Českým rozhlasem, jestli o tom můžete mluvit.

PK: Super, tak se pusťme do toho. My jsme se ptali, jestli vám vůbec něco můžeme říct a bylo nám řečeno, že je to za Český rozhlas OK. Rámcově tam máme nějaké scénáře. My samozřejmě máme stav as it is a pak máme stav, který bychom chtěli v budoucnu. Společně jsme měli několik projektových schůzek v Českém rozhlasu, kde jsme si vydefinovali i budoucnost, která je poměrně vzdálená od toho, kde to je teď. A to ne tak z naší strany, protože my to máme vlastně připravené, ale spíš ze strany Českého rozhlasu, kde je potřeba úprava toho produktu tak, aby to bylo schopné využít potenciál, který náš produkt nabízí.

TL: Obecně to doporučování, které je na Českém rozhlasu, funguje na kombinaci více doporučování. Je to doporučování vyloženě na uživatele, což znamená, že když uživatel přijde na stránku, tak mu podle konkrétní historie doporučí nějaký obsah. Druhým typem doporučování je, když doporučujete vhodné položky k nějaké položce, ať už je to třeba audio, článek nebo něco podobného. Český rozhlas má oba tyto typy doporučování. Budu říkat Český rozhlas a mluvím tím o iRozhlase i o aplikaci mujRozhlas. Takže Český rozhlas má oba tyto typy doporučení. Standardně fungují na nejrůznější kombinaci našich algoritmů, ať už je to něco na způsob kolaborativního filtrování, což znamená, že doporučujeme nějaké itemy podle toho, jak byly interagovány s uživateli, kteří jsou podobní tomu konkrétnímu uživateli. Kombinujeme tam podobnosti těch itemů, které odvodila AI na základě embeddingů, které k tomu itemu jsou. Zahrnujeme tam popularitu, což znamená, že vyhodnocujeme, jakým způsobem je konkrétní položka v tuhle chvíli populární a trendová na daném konkrétním místě atd. Obecně ty modely fungují takto. Pro každý máme scénář, a je to jakoby doporučovací místo, což znamená, že třeba na stránce k pořadu máme například *následující pořady* nebo *poslechněte si také*. Každý ten scénář je řízen nějakou logikou upravenou podle naší best practices nebo nějakých specifických požadavků Českého rozhlasu. Potom je řízena ještě nějakými dalšími pravidly, což znamená, že například můžeme manuálně vyloučit nějaký pořad z doporučení, nebo naopak můžeme podpořit nějaký pořad, který je v tu chvíli editorskou volbou. Takže se snažíme zároveň neomezovat vliv editorů a chápeme, že i jejich pohled je velmi důležitý. No a každý tenhle scénář tak má samozřejmě za sebou nějaké AB testování, kdy jsme nastavili více variant a vyhodnocujeme, co funguje nejlépe. To je úplně tak jako k začátku, jakým způsobem jsou ty scénáře nastavené. Pak také funguje i nějaká vzájemná korektivní zpětná vazba, kdy se díváme na výkonnost. Nemáme takovou tu novinářskou odbornost, nemáme produktové cítění a ani tu

produktovou odpovědnost za ten konkrétní výsledek, takže nejsme schopni říct, jestli to doporučení z hlediska novinářského nebo editorského vyhovuje tomu, co by si ti editoři na těch konkrétních místech představovali. Takže to funguje na nějakém principu zpětné vazby, kde pak třeba to doporučování ladíme. Nebo v průběhu, když nám Český rozhlas řekne „proč tady, já vidím tohle a tohle doporučení,“ tak my vyhodnotíme, proč se to děje, potvrdíme si, že je to správně nebo případně upravíme.

A jak to probíhalo? To si dáte schůzku s těmi editory nebo s těmi manažery?

TL: Pro nás nejsou primární ti editoři, nebo nejsou zahrnuti v tom procesu, ale řešíme to spíš přes ty produktové lidi, což je pro nás vlastně Honza Misák a Dalia Moudrá v tuhle chvíli. Na webu máme i sekci AI v médiích, která je pro čtení poměrně uživatelsky přívětivá. Samozřejmě vnímáme i nějakou zodpovědnost a snažíme se být nápomocni. Třeba pod Pavlovým vedením vznikl open source, který vyhodnocuje doporučování.

PK: Já bych řekl, že ty algoritmy už máme tušený jinak, než jenom ve smyslu, že „tady potřebujeme kliky a uzavíráme člověka do bubliny.“ Vlastně už od začátku, a to byla právě třeba disertační práce jednoho našeho spoluzakladatele, jsme se zaměřovali na to, jak naopak pomoci uživatelům najít obsah, který je mimo jejich bublinu. Takže to už je zakódováno v základu toho našeho řešení. Chceme ukázat, že vyhodnocovat ty rekomendry jenom pomocí kliků je hloupost. Právě proto jsme začali zveřejňovat open source software a máme nějakou komunitu v rekomenačních systémech, která se zabývá právě tím, jak vyhodnocovat ty doporučovací systémy, aby bylo vidět, co to s těmi uživateli dělá.

Jak se liší ten produkt pro zpravodajské médium, jako je iRohlas a jak se liší pro komerční subjekt?

PK: Vždycky záleží na tom, jaké nám médium zadá KPI (klíčové ukazatele výkonnosti) a performance indicators, co má ten recommender optimalizovat. Když je to nějaká služba, která má subscription, tak je to poměrně jednoduché, protože tam se optimalizuje spokojenost uživatelů, aby prodlužovali předplatné. Potom jsou média, kde to je trochu složitější, protože například vydělávají na reklamě a musí optimalizovat trochu jiné kritérium než jenom spokojenost uživatelů. Tam jsou doporučovací systémy nastavené tak, aby jim to, když si řeknou, přinášelo rozumný profit i třeba z reklamy a tak dále. U Českého rozhlasu je to jednodušší v tom, že jsou veřejnoprávní médium, takže tam není potřeba jakoby honit nějaké extra page views a boostovat třeba v tom recommenderu obsah, který by normálně uživatelé nezhlédli, jenom kvůli tomu, že je potřeba prodat nějakou reklamu. Na druhou stranu proto, že veřejnoprávní médium má nabídnout vyvážené zpravodajství a tak dále, tak se nějakým způsobem asi do budoucna uvažuje o tom, že by se právě boostovaly různé typy obsahů. Můžete si to představit jednoduše tak, že třeba každý editor v té redakci bude mít nějaké hlasy. Editor řekne, „tohle jsou nějaké moje oblíbené články v poslední době, které si myslím, že by měly plnit tu veřejnou funkci líp.“ Tím pádem oni si můžou nastavit, aby to ten recommender bral v potaz, nezobrazoval to úplně všem, ale jenom zvyšoval pravděpodobnost, že to ti uživatelé uvidí. Tím v podstatě můžeme zvyšovat sledovanost článku jakoby méně invazivním způsobem, které by třeba jinak nebyly tak populární. Snažíme se, aby to bylo vyvážené. Média si to nějakým způsobem chtějí hlídat a my tomu jdeme naproti. **Důležitá je role editorů, proto** teď zveřejňujeme postupně novou funkcionalitu Insights, která umožní provozovatelům trochu lepší vhled do toho, co recommender dělá. Například si můžou v reálném čase sledovat nějaký ten doporučovací scénář, a podle toho můžou ovlivňovat doporučování tak, aby to respektovalo jejich cíle. Myslím si, že se posouváme z toho „tady běží black box recommender, který to nějak udělá“

do pozice, kdy skupina projektáku nebo editorů má poměrně dobrou kontrolu nad tím, co se tam děje.

Můžu si to představit jako nějaký program?

PK: Je to v podstatě panel editora, který má takový jako kokpit, kde řídí ten rekomenční engine.

TL: Jak jsme říkali, na jednotlivých doporučovaných místech můžeme mít každou metriku jinou. Což znamená, když máte nějaké velké médium, které v sobě kombinuje různé typy obsahu, ať už je to ryze zpravodajský, volnočasový atd., tak každý ten scénář se může chovat jinak. U těch zpravodajských se snažíme dávat poměrně velkou míru volnosti editorům, kteří mají „*editorsko-produktový feeling*“, mají představu o tom, kam to médium směřovat, takže mají poměrně velkou možnost ovlivnění doporučovací plochy.

U komerčních médií může být jednou z hlavních metrik zvýšení počtu předplatitelů, zvýšení počtu prokliků, prodloužení průměrného času stráveného na stránkách, dočtenost apod. Tedy všechno, co zvyšuje příjmy, ať už z předplatného nebo zobrazení reklamy. U veřejnoprávního média jsou to také samozřejmě důležité metriky, protože jsou indikátorem uživatelské spokojenosti, ale do hry vstupuje i informační role a snaha doporučit uživatelům vyvážený zpravodajský obsah, i na úkor například proklikovosti. Nastavení systému je tedy jiné. Co se týče uzavírání do bublin, jsme poměrně flexibilní v tom, že si tam umíme nastavit diverzifikaci, kterou ty články mezi sebou mají mít, abychom člověku, který se třeba zajímá v té zpravodajských věci o politiku a například konkrétně o Kongo, nevytápeli stránku Kongem, byť tam ty články jsou, ale abychom mu opravdu poskytli vyvážené zpravodajství. Máme i takzvané *must have články*, které se v tom doporučování musí objevit. Kombinujeme tam opravdu aktuální trendové články a breaking news, které se tam objevují. Jsou to ty články s editorskou váhou plus mix, opravdu toho personalizovaného doporučení Konga. Zatímco když je to nějaká volnočasová stránka v tom samém médiu (třeba stránka *zajímavosti*), tak mu tam to Kongo klidně dáme, pokud ho to opravdu zajímá s tím, že ho občas zkusíme vylákat z bubliny. Ale opravdu každé to konkrétní místo se liší a snažíme se editorům dát volnější ruku i tím, že rozhraní, které mohou používat, je poměrně uživatelsky přívětivé a nevyžaduje právě nějaký program. Je to opravdu o nějakém jednoduchém manuálním nastavení.

Takže, když tohle zavedete v nějakém médiu, tak musíte proškolit ty editory, aby viděli, jak to funguje?

TL: Počítáme třeba s tím, že poskytneme nějaké úvodní školení zaměstnancům a je potom otázkou té druhé strany, jestli to bude školení pro ty produktáky, kteří to potom předají dál, nebo jestli to bude přímo pro editory. My bychom tam ty editory, rádi zachytili víc. V tuhle chvíli jsou pro nás primárně těmi komunikačními partnery produktáři a programátoři.

PK: Zákazníci k tomu mají různý přístup, někde to udělají tak, že mají nějaké svoje CMS, kde to editují. Přidají jim například nějaký checkbox „tohle je my favorite“ a oni technicky nerozumí tomu, co se tam děje, že to jde do recommendru, a tam se to boostuje. Otázka je, jak moc tomu musí rozumět. Pro ně je důležité, aby zvýšili impakt a myslím si, že například sledování toho, jak vypadá audience, je spíš pro ty projektáky.

Prostě pochopit to, není úplně jednoduché. Když si vyjedeme nějaká data z toho recommendru a díváme se, jak se ty lidi chovají atd. a analyzujeme to, tak je to hodně advanced tool. Je to těžké jednoduše vysvětlit lidem i z médií, kteří nemají vlastně ten technický background. A například neví, že tyto osy, nemají fyzický význam, že to je jenom projekce do latentního prostoru a jak to vlastně celé funguje.

TL: Já si myslím, že se snažíme, že na blogu máme nějaké edukativní články, a ještě to plánujeme trochu podpořit.

No to by mě taky zajímalo, jak to funguje.

PK: Jestli vás to zajímá, tak my máme na FITu na to bakalářský, magisterský program a je dobré začít na začátku nějakou tou algoritmizací, potom člověk potřebuje statistiku, aspoň nějakou lineární algebru a další věci, aby chápal práce s maticemi... když tomu chcete jako vnitřně rozumět.

TL: No já si myslím, že ne. Opravdu pro ty editory je to samovysvětlitelné. Nemusíte vědět, že je to počítáno nějakou kosínovou podobností, musíte vědět, že když je to blízko, tak to znamená, že si jsou ta data podobná. A jak říkám, myslím, že jsme rozjeli poměrně hezkou sekci na našem blogu a chystáme se to vysvětlit do větší míry. Není důvod se toho nějak bát, nebo mít dojem, že k tomu musím mít technické vzdělání, to určitě ne. Tam to jenom chce věřit nám, že ta data jsou reálná a že ty tečky tam nejsou náhodně.

PK: Těch algoritmů jsou stovky a jsou různého druhu... Spousty toho najdete v těch blogpostech.

Editoři ještě zmiňovali, teda nemám zatím rozhovory se všemi, ale setkala jsem se i s nějakou jejich potřebou ohledně tagování v rámci redakčního systému, kde třeba zmiňovali, že je to pro ně práce s doporučením v tom smyslu, že jim tam vyskočí nějaké tagy, které si oni můžou zaklikat. Jenže ty tagy nejsou natolik chytré.

PK: To jsou věci, které se týkají hlavně automatického tagování obsahu. Tam je taky AI, ale je to trochu jiný druh AI než doporučovací systémy. To můžou být právě jazykové modely nebo další věci, které z textu extrahují nějaké ty tagy a metadata. To se učí jiným způsobem a na jiných datech než doporučovací algoritmy... ale je jasné, že když je to dobře otagované, tak to ten doporučovací algoritmus může vzít v potaz. My si třeba tady tohle děláme interně a ty externí tagy vlastně moc nepotřebujeme, protože ty naše doporučovací algoritmy jsou tak advanced, že tomu obsahu dokáží porozumět.

TL: My ani nepotřebujeme ty tagy. My rovnou děláme embeddingy, které pak využíváme pro to doporučení. Tagy jsou pro nás vlastně mezifáze, která pro nás není důležitá, když to takhle řeknu, takže třeba tohle pro Český rozhlas neděláme my a nevím jakým způsobem to mají implementováno.

PK: No jde to trochu směrem k vysvětlitelnosti... to je taky další téma, které se může probírat. Vysvětlitelnost toho, proč to doporučil ten doporučovací systém pro toho uživatele. Tam se ta veřejnoprávní média třeba snaží, aby ti uživatelé měli víc vhledu, takže aby tam bylo „toto bylo doporučeno strojem a bylo vám to doporučeno z toho a z toho důvodu.“ Nicméně se jednak ukázalo, že ty uživatele to zas až tak moc nezajímá a druhá věc je, že je to hodně těžké udělat jako fakt upřímně. Protože když máme 100 různých algoritmů, které vstupují do toho, jak ta rekomendace může vypadat a zároveň jsou tam nějaké editorské boosty, které ještě tu rekomendaci dál ovlivňují, tak napsat jednoduše „proč bylo doporučeno zrovna tohle“ není v lidských silách. Takže pak se používají nějaké trochu proxy modely na tu vysvětlitelnost. Například že se řekne: „to by asi mohlo vycházet z toho, že si předtím interagoval taky s tímto typem obsahu.“ Ale je to hodně hrubé přiblížení toho, proč to bylo doporučeno.

TL: Z naší strany mi jako nepřijde, že by to bylo téma. Myslíte, že to je téma?

Jako jestli to tam uvádět?

TL: Jako ta vysvětlitelnost, jestli to tam uvádět z hlediska uživatele.

To je taky jedna z mých otázek na vás. Nevím jak u uživatele, ale mezi novináři se setkávám s tím, že to tam chtějí mít. Myslím si, že pro toho uživatele to asi není až tak důležité, ale pro ten mediální dům ano, protože jde v nějakém smyslu o transparentnost. Jaký je váš názor na to?

PK: My se snažíme ten produkt rozvíjet podle potřeb našich zákazníků a oni mají stovky zákazníků z celého světa a mají nějaké feature requests. My jejich požadavky vždycky dáváme do *to do listu* na hromadu a snažíme se odbavovat ty, co zasáhnou nejvíce zákazníků. A tohle se nám v tom seznamu nikdy nedostalo nikam výš. Nevím proč, ale není o to prostě takový zájem.

TL: Není to téma. Velmi zřídka se stává, že od nějakých technical advanced lidí přijde prosba o to, abychom třeba řekli, jaký konkrétní model doporučil, ale to je velmi sporadicky. Nesetkali jsme se s tím, že by naši uživatelé, ať už to jsou čtenáři nebo poptávají cokoli z hlediska produktu, řekli „chceme rozumět tomu proč tohle konkrétní doporučení.“ Jak říkám, může se stát, že někdo má dojem, že to doporučení neodpovídá tomu, co by čekal. Pak samozřejmě uděláme revizi. Může to být třeba typický příklad u těch mediálních domů, kde se stává, že to doporučení vypadá jinak pro normálního čtenáře a pro editora, protože ten editor má samozřejmě tu historii mnohem širší, a to chování je mnohem posunutější. V tom případě dokazujeme, že to jeho doporučení je spíše raritou než tím, co se běžně stává a řešíme to třeba testovacím účtem, kdy si ověříme, že to doporučování běží opravdu kvalitně. Ale, jak říká Pavel, to že bychom čelili nějakým požadavkům o vysvětlitelnost, tak to v tuhle chvíli tak opravdu není.

A jaký je váš názor? Mělo by to tam být napsané někde? Měl by ten čtenář vědět o tom, že mu je něco doporučováno automaticky?

PK: Já si myslím, že ne. Protože vlastně to je nový standard.

TL: Přesně tak.

PK: Protože lidi jsou zvyklí na to, že jim je něco doporučováno automaticky už skoro všude. Facebook, Twitter, prostě kamkoli kouknou, tak je to doporučováno automaticky. Takže by tam mělo spíš být „Ne my jsme 100 let za opicemi, a ještě vám to nedoporučujeme automaticky.“ Že ta budoucnost je jako jinde.

TL: Já si myslím, že lidi si opravdu neuvědomují, kolika s prominutím balastem by se museli brodit, kdyby tam to doporučování nebylo, protože toho obsahu opravdu vzniká strašně hodně a ten doporučovací systém je takový třídič pro to, najít uživateli relevantní informace. Takže ten život by byl mnohem mnohem horší bez doručovacích systémů.

PK: Taková ochrana trošku. Druhá věc je, že někteří uživatelé by si to rádi customizovali, aby to doporučování bylo víc podle jejich představ a aby měli víc pod kontrolou třeba historický uživatelský profil a tak dále. Z našich zkušeností je to strašně malé procento uživatelů a je otázka, zda se vyplatí investovat do vytváření user interface pro konfiguraci. Bylo by to hezké, ale produktově to příliš nedává smysl a je otázkou, zda by to bylo využíváno. Bylo by to hezké, ale produktově to příliš nedává smysl a je otázkou, zda by to bylo využíváno.

TL: Mám zkušenost z nějakých médií, že tam, kde se ta možnost čtenáři udělala, ne teda na úrovni přímo toho doporučení, ale na možnosti ovlivnit si, jakým způsobem má to zobrazení vypadat, tak to procento lidí, kteří to objektivně využívali, bylo tak mizivé, že si myslím, že to opravdu není nějaký problém. A jak říkám, pro nás opravdu není dramatický problém vytvořit ke každému doporučení nějakou větu, která teoreticky toho člověka může uspokojit, ale tu přidanou hodnotu tam osobně nevidím.

ČÁST II Táňa Lančová

Kdyby teoreticky vznikla stránka, která by byla takhle personalizovaná, budete potřebovat, aby si tam uživatel vytvořil účet? Na základě jakých dat to doporučujete?

Fungujeme na bázi first-party, což znamená, že vyhodnocujeme chování uživatele pouze na stránkách, kde doporučujeme a s obsahem, který tam je. To znamená, že jestli si právě koupíte boty, vzápětí Vám nedoporučíme článek o botách, protože tuto informace nemáme a nechceme.

Pro menší zákazníky máme připraven skript, který si mohou přidat do stránky a který nám kliky uživatelů zasílá. Pro větší mediální domy je ale standardem, že se integrují takzvané backendově, to znamená že mají identifikaci uživatelů i sběr dat řešen na své straně a do rekorderu je jen zasílají. Jakým způsobem to dělají, je pak jen výhradně jejich rozhodnutí. Může to být přihlášený uživatel, cookie, session nebo mix všeho. Nejlepší je samozřejmě přihlášený uživatel, protože pro něj dokážeme připravit personalizovaný zážitek skrz různé platformy. Ale dokážeme ho zkvalitnit i pro uživatele definované cookie, dokud si ji nesmažou, nebo i jen v rámci jedné session.

Takže je to provozovatel, kdo si to musí technicky nějak ošetřit nebo zpracovat?

U těch velkých ano. Je to standard, že má řešenou identifikaci uživatele. Pro ty malé poskytujeme nějaký skript, který si můžou umístit na svoji stránku, ale stejně jako u těch velkých mediálních domů je standardem, že to mají už dávno řešeno sami.

Ještě by mě zajímalo, jestli jsou podle vás novináři nebo editoři všeobecně připravení na to s tímhle pracovat? Nebo jak už tady částečně zaznělo, jestli je to pro ně moc složité nebo není?

Já si nemyslím, že to je složité, já tam vnímám takový velký gap, který se nám nějak nedaří překlenout a myslím si, že to je i v těch samotných redakcích, že je velká propast mezi těmi produktovými manažery a editory. Myslím si, že to AI je opravdu velmi významný pomocník, který opravdu ulehčí uživateli cestu k obsahu. I z hlediska editorů. Jim opravdu pomůže doporučovat ten obsah, který si myslí, že je podstatný pro relativní část uživatelů a zároveň dá příležitost longtailovým článkům, které by normálně v běžném nepersonalizovaném provozu mohly zapadnout. Oni se tam protočí a ten recomander je schopný ho vytáhnout tomu konkrétnímu uživateli.

Když bychom to srhnutí, jaké ty výhody a etické výzvy by to přineslo z pohledu čtenářů a konzumentů takového obsahu a z pohledu těch editorů?

Tak z hlediska uživatelů je to snadnější cesta k relevantnímu obsahu, protože toho obsahu se opravdu můžou produkovat stovky denně a recomander tu cestu k tomu velmi zkrátí. Dobře udělaný recomander je udělaný tak, aby tam měl nějaké jisté procento diverzifikace, což znamená, že vás neuzavře do té bubliny. Zároveň reaguje pružně na každou interakci, kterou uděláte, takže je schopen zachytit aktuální zájmy. Také vyhodnocuje interakce všech uživatelů, podchytí tedy trendující články. A pokud je skutečně dobrý, umožní i editorské nastavení, tedy uživatele neochudí o skutečně důležité informace, které sice neodpovídají jeho zájmů ani nejsou populární, nicméně by neměli utéct ničí pozornosti. Takže z hlediska uživatele si myslím, že tam jsou samé výhody.

Z hlediska editorů si myslím, že to může být opravdu velký pomocník, když si na to doporučování najdou „ne black boxového“ partnera, ale partnera, který je schopný poskytnout jim ad jedna to iniciační nastavení dle jejich přesných požadavků a ad dva

umožní mít vhléd do těch dat, do doporučení, do chování uživatelů a ad tři jim umožní ta doporučení nějakým způsobem ovlivňovat.

Ještě mi v těch rozhovorech zazněla dejme tomu nějaká obava o to, jestli ta data, která se sbírají, nejsou příliš citlivá. Například tam byl zmíněn případ Netflixu, který začal doporučovat lidem, o kterých zjistil, že jsou třeba queer, hodně podobného obsahu. I když předpokládám, že ti uživatelé s tím doporučením jsou spokojení, zajímala by mě ta debata, jestli tohle zrovna není moc citlivý údaj například?

Já si myslím, že ale Netflix nefunguje na tom, že by měl u toho uživatele nějakou poznámku queer a podle toho doporučoval, ale že doporučuje opravdu podle chování. Takže v momentě, kdy ten queer uživatel začíná více klikat na dětské animáky, tak mu zase bude doporučovat dětské animáky. Je to takový častý problém, že je to ten marketingový pohled na to doporučování, který je asi přístupnější a je takový jako známější... takže si řeknete, fajn, tak tady máme skupinu lidí, kteří se zajímají o zprávy, tady máme skupinu lidí, kteří se zajímají o auta a tady máme třeba ženy 25 až 30, které zajímají nákupy nebo něco takového a každému uděláme profil, kam ten uživatel spadne a podle toho mu doporučujeme. Nějaký basic recommender takhle fungovat může, ale opravdu dobrý recommender funguje na tom, že doporučuje unikátně každému uživateli. Takže nám je úplně jedno, jestli máme uživatelské ID a je nám úplně jedno jestli to uživatelské ID je muž, žena atd. My prostě doporučujeme na základě uživatelského chování toho uživatele a na základě těch pravidel, na kterých jsme se s tím editorským týmem dohodli, nebo podle aktuálního editorského nastavení, ale neděláme si nějakou takovouhle metaanalýzu. Samozřejmě nám můžou na speciální přání tato uživatelská data poslat a funguje to třeba, když doporučujeme taky v médiích, ale když doporučujeme ve více zemích, například jako nějaká streamovací platforma, která působí ve více zemích, tak tam pak je třeba cennou uživatelskou informací preferovaný jazyk uživatele, kdy můžeme doporučit obsah, který je v tom jazyku, který si tam uživatel zadal.

Může si uživatel resetovat to nastavení?

Záleží na tom, jakým způsobem je to ID toho uživatele, jestli to je na cookies, jestli to je na heslo atd. Ale nějakou možnost typu „teď mi smaž moji historii,“ tak tu máme ve smyslu, že nám ten zákazník, ne ten uživatel, ale ten zákazník může kompletně smazat historii uživatele. To samozřejmě poskytujeme a je to pak na základě toho, jak to má řešeno. Vím, že když řeknete „smažte mi celou moji historii“, tak je vám povinen vyhovět a k nám se to dostane takhle zprostředkovaně, ale že bychom měli něco jako magic a „totálně mi změň doporučení,“ tak na to dosud nebyl nějaký požadavek. Ono to doporučování je málokdy na sto procentech obsahu. Vždycky ten doporučovač pustí na určité to doporučovací místo, takže uživatel má opravdu mnoho příležitostí dostat se i k jinému obsahu. I kdybychom ho do té bubliny teoreticky uzavřeli, jako že to neděláme. Opravdu se to doporučení mění s každým klikem, s nově vydaným obsahem atd. Není to o tom, že byste si třikrát klikla na nějaký článek třeba o Britney Spears nebo něco takového a my bychom vám pořád dávali jenom nějaké celebrity atd. Tak takhle to nefunguje. Ta doporučení nejsou statická, jsou diverzifikovaná a pružně reagují na chování uživatele. Pokud ale myslíte skutečně nastavení uživatele ve smyslu třeba preferovaný jazyk, kategorie atd., které si uživatel dělá sám, pak by mu rozhodně měl tuto změnu poskytovatel umožnit.

Ještě by mě zajímalo ohledně toho uvádění nebo té transparentnosti a označování toho, že je něco doporučováno automaticky. Zmiňovali jste to, že pro vás je to nový standard

a že se to neuvádí, protože všichni počítáme s tím, že je nám něco automaticky doporučováno. Ale přijde mi, že legislativa teď jde tím směrem, že vyžaduje po společnostech jako Facebook nebo TikTok apod., aby byly více transparentní ohledně toho doporučování. Jaká je ta situace u nás a jaký je váš pohled na tohle?

Opravdu věříme, že doporučování pomáhá a zaručuje standardní chování webu nebo aplikace, které a priority jako konzumenti obsahu očekáváme. Jen je třeba trochu větší edukace o tom, jak funguje a z jakých dat vychází. U nás se opravdu nesetkáváme s tím, že bychom měli požadavky na vysvětlení. Když se zákazníkům zdá, že by to doporučení mohlo vypadat lépe, tak dostaneme dotaz, proč to vypadá takhle, a vysvětlíme to. Ale nějaký systematický požadavek o tom, abychom zdůvodnili doporučení, v tuhle chvíli opravdu nemáme. Opravdu si myslím, že by byli lidi překvapení, kdyby se u těch služeb, kde jsou na ta doporučení zvyklí, lusknutím prstu vypnula. Myslím že, by to opravdu bylo velké překvapení, jak to vlastně funguje a neodpovídá.

Respondent Čeněk Třeček, datum 25.10. 2023

Víte, co jsou to rekomenační algoritmy? A jak fungují?

Ne.

Vzpomínáte si na řádek Robot doporučuje na iRozhlase?

Jo, jo.

Co si o něm myslíte?

Vlastně jsem ho nikdy moc nepoužil.

Pracujete teď nějak s automatickým doporučováním v rámci vaší editorské práce?

Pouze vlastně u klíčových slov toho textu. Ale je pravda, že já to většinou stejně dělám víc podle sebe, než že bych si to nechal... Občas se schválně podívám, co mi nabídne, ale stejně se rozhoduji víc podle sebe. Takže podle toho, co mi tam dá.

A podle jakých třeba kritérií se teda rozhodujete?

Podle toho, co je v tom textu. On si vybere ten algoritmus třeba hodně místní jména nebo tak a myslím si, že třeba nejsou tak důležitý jako ta událost, která se děje a kterou on třeba nevyhodnotí.

Kdybychom si představili teoreticky, že by vznikla nová stránka iRozhlasu, která by byla čtenáři ušitá na míru jako nějaká záložka, nebo druhá verze té stránky.... co byste na to řekl?

Myslím, že se o tom už dlouho mluvilo, nevím, jestli to někdy vlastně vešlo v platnost, ale že podle toho, co by se hledalo, by se i zobrazovalo tomu člověku. Vlastně jsem to nikdy v praxi neviděl. Jako dovedu si to představit, myslím, že to jde udělat, ale zatím to nemáme.

Máte na to nějaký svůj názor, jestli je to potřeba...

Já si myslím, že to jako určitá cesta je, že prostě ty čtenáře se tam takhle individuálně dají udržet nebo vlastně získat tím, že tam opravdu najdou to, co je pro ně zajímavé. Ale zas to teda svádí k tomu, že se ten člověk dostane do určitý bubliny. Že vlastně některý názory prostě ten algoritmus nevyhodnotí a k němu se vůbec nedostanou. Je to vlastně to samé, jako na sociálních sítích.

Tím jste vlastně už částečně zodpověděl další otázku, jaké výhody nebo etické výzvy by to přineslo... Co by to podle Vás přineslo vám, jako editorům? Jestli by to třeba změnilo nějak vaší práci, nebo vůbec ne?

Jako umělá inteligence a její využití teď? Asi jo. Já si myslím, že časem ji změní, ale v tuhle chvíli vlastně to jako je v plenkách v podstatě. No, tuhle jsme řešili, že by nějakým způsobem by nahazovala zprávy z Četky. To by bylo jednodušší, ale otázka je, kdo bude rychlejší. Já jsem celkem rychlý. To jsou vlastně takový jako opravdu technický věci, doplňování nějakých asetů, ale co se týká toho textu, té zprávy, tak si myslím, že to AI úplně nikdy nemůže nahradit a ani to není žádoucí. Protože ta zpráva je opravdu něco, co se udělo, když se to čte a ne, že to vytvořila umělá inteligence. Ono to souvisí potom i s fotkami, na které narážíme, protože náš fotograf by chtěl dělat ilustrační fotky pomocí umělé inteligence a zatím se tomu bráníme. Protože vlastně to potom bude vypadat divně, když je tam zpráva,

nějaké poctivé tvrzení, že se někde něco stalo, nějaká událost, a pak tam bude něco, co je vlastně umělé.

A konkrétně ty doporučovací algoritmy, jestli by to nějak změnilo tu vaši práci nebo ne... To, že by vlastně něco za vás vybíralo nebo doporučovalo články?

Já jsem se s tím setkal, ale ne na iRozhlase. Ještě na iDNESu jsme měli algoritmy, kdy nám to vlastně ukazovalo (ono to teď tady taky je, ale nikdo to moc nepoužívá), jak si to vede ten článek v pořadí na tom místě, kde je na té webový stránce. Jestli je to tak optimální, jestli třeba nemá vyšší potenciál, nemá se posunout výš nebo naopak, že o něj není zájem a měl by být níž. Takže tohle už existuje, to už se dělalo, problém je ten, že tam jsou věci, které si myslím, že jsou dobré, takže by tam měly být, i když vím, že ten zájem o ně nebude tak velký. Není to zase click bait, když to tak řeknu, ale jde o důležitou zprávu a toto zase ten algoritmus nevyhodnotí. Ještě se na iDNESu řešily titulky *a* a *b*, kde se zvolily dva titulky, částí lidí se ukazoval titulek *a* a částí čtenářů titulek *b* a vlastně se vyhodnocovalo, který si líp vede a potom se dalo podle toho určit, který nakonec bude vydán.

Takže vy jste navrhly třeba dva titulky...

...třeba se navrhly dva titulky a podle zájmu čtenářů, jejich úspěšnosti u těch čtenářů, po nějaké době, se tam vlastně už držel jen ten jeden, ten úspěšnější.

A to, jak jste o tom mluvil předtím... jak to vypadalo? To bylo nějaké rozhraní...?

.... to byl jakoby semafor v rámci těch pozicí. Když jsem vlastně viděl na homepage ty články pod sebou, tak tam byl semafor, kde se ukazovaly barvy žlutá, červená, zelená. Zelená byla v pořádku, ukazovala na nějaký potenciál článku, oranžová byla ještě v normě, tak akorát. Ta červená tlačila články někam dolů, níž. A u nás tady na iRozhlase přes Chartbeat to jde taky, když si představíte web iRozhlasu, tak u každého článku bylo vlastně pořadí čtenosti. Takže tohle je čtené víc, tohle míň, a taky se dalo poznat, o které články je zájem a například když jsou nízko, a přitom o nich zájem je, tak že by se daly vytáhnout.

A na tom iDnesu šlo hlavně o klikanost?

U toho iDNESu šlo čistě o čtenost. A tady na tom Chartbeat, to je to samé, počítá ty přístupy.

Stalo se vám třeba, že svítil červeně nějaký článek, který Vy byste jako editor označil zeleně?

No, jo určitě se stalo, že svítilo jinak něco, o čem jsem si myslel, že je to důležitá zpráva.

Takže dalo by se říct, že vám to práci usnadnilo?

Tehdy trochu jo. A i teď se dá na to podívat, ale jak říkám, je to dvojsečný...

V čem?

Je to dvousečný v tom, že pak vy pojedete jenom na čtenost a vlastně ztrácíte tvář. Každé to médium je nějakým způsobem postavené, dává nějakým zprávám nějakou důležitost a tohle to z toho ztratíte. Prostě dojdete na to, že se budete jenom podbízet, dostanete se jenom do jednoho žánru.

Jaké jsou třeba teda principy nebo hodnoty, které vy máte teď nastavené, podle čeho vy zprávy vybíráte.

No, to je otázka... Myslím si, že za prvé podle toho, jaký má ta zpráva dosah, dopad. Hodně řeším, jestli to je naše zpráva nebo není. Jestli to tam udržím, jestli jsme to psali my, jestli jsme si s tím dali tu práci. Určitě na prvním místě je to, jestli to je naše zjištění a jestli je to opravdu zpráva, o které se bude mluvit, která má šanci, že ji někdo od nás přebera. To jsou ty zprávy, které se snažíme dostat nejvýš. No a co se týká toho, co dám do té naší desítky, kterou tam máme, tak aby to bylo trochu pestré, aby tam byla například i nějaká kultura. Samozřejmě jsou to zprávy, které mě zaujmou, asi je tam i to subjektivní hledisko. Mám to třeba nastavené i jinak. Na iDNESu, protože dřív jsem to dělal jinak. Navíc jsem dělal černou kroniku, takže jsem volil zprávy trochu jiného typu než teď. Naopak tu černou kroniku mám teď hodně upozaděnou, protože vím, že jsem to dělal, abych právě do toho nespádl, a netahal ty věci, o kterých bych věděl, že se budou číst, ale teď je tam nechci. Jsme veřejnoprávní...

V komerčních médiích se díky rekomendačním algoritmům snaží zvýšit personalizaci, tak jestli má používání rekomendačních algoritmů opodstatnění v Českém rozhlasu ve veřejnoprávním médiu. Proč ano a proč ne.

Asi jo, já jsem říkal, že je fajn, že se to tomu člověku trochu přizpůsobí, ale nevím, jestli by neměl mít člověk možnost si to sám ovlivnit. Jako že si řekne, že tohle chce, tohle ne. To už není potom algoritmus umělé inteligence nebo trochu jo. Nebo aspoň aby člověk věděl, že se něco takového děje. Rozhodně by to mělo být přiznané v tom případě, že by se to dělalo. Vlastně se tomu úplně nebráním.

A jak byste to vyřešil? Že byste to k článku napsal?

Nějakou poznámkou nebo nějakým úvodním nastavením. Ale pro každého, kdo to začne využívat, by tam mělo být upozornění, že takovou věc používáme. Protože to přesně svede k tomu, že to prostě skončí někde v nějaké bublině.

A v tom iDNESu jste to označovali nějak?

Tohle jsme neoznačovali, to tam nikdy nebylo. Jak jsem říkal, to řazení článků byla vlastně informace pro toho editora, a to se opravdu řídilo podle čtenosti. Tam to tenkrát nebylo individuální tak, že by na člověka vypadly ty články. Ty články byly tak, jak je seřadil editor. Nebylo to tak, jak by to algoritmus seřadil pro člověka podle toho, co ho zajímá. Tam vidím ten rozdíl.

Takže to vlastně nebyla úplně umělá inteligence.

Vlastně ne, je to jako nějaký pomocník toho editora, který jede přes tu čtenost, ale není to tak, že by umělá inteligence říkala, že tomuhle čtenáři se dá tohle a tohle.

Takže spíš nějaké vaše vodítko...

Takové vodítko, přesně tak.

A když jste tam s tím pracovali, tak vás na to nějak zaškolili?

Vlastně nám to jednou ukázali, ono to bylo celkem jednoduchý, takže na tom nebylo nic složitějšího.

A jinak, jestli se teďka u nás něčeho týká umělá inteligence, tak to budou přepisy.

Měl byste zájem o nějaké školení nebo manuál od projektových manažerů, kdyby se něco takové zavedlo?

Upřímně, nic jiného by mi nezbyvalo. Pokud bych s tím chtěl pracovat, tak bych to samozřejmě musel ovládat. Ale jinak ot asi nepotřebuju.

Používáte při práci nástroje jako například ChatGPT, Whisper nebo Midjourney? A pokud ne, tak proč je nepoužíváte?

Neznám je, a vlastně tím pádem s nimi ani neumím pracovat. Nenašel jsem nic, k čemu bych je nutně potřeboval.

A zkoušel jste si je třeba?

Ne.

A třeba redaktori s tím pracujou nebo nějak?

Předpokládám, že směnaři, kteří pracují na denní provoz určitě ne a jestli to zkoušel někdo z redaktorů, tak to opravdu nevím.

A například přepisy mluveného slova do textu, o kterých jste se už zmiňoval, používáte?

Přepisy používáme, ale stejně se musí kontrolovat. Aktuálně máme v našem systému oficiálně dva přepisovače. Jeden přepisuje pomalu, ten, který dělal Cíba (Jan Cibulka) přepisuje o něco rychleji, ale vždycky se to musí s tím zvukem sladit. Například bývá problémem přepis jmen, které se nikdy nepřepíší úplně přesně, zvláště cizí jména. Zatím tuto funkci občas používáme, protože stejně zvuk z vysílání přepisujeme, a pokud se jedná o dlouhý záznam, tak ten přepis aspoň trochu pomůže práci usnadnit. Anebo to pomáhá například při rychlé orientaci, když nahlédnu do přepisu určitého vysílání, tak aspoň zpětně tuším, o čem se tam tak bavili.

A naopak co říkáte na to, pokud by články na iRozhlasu četl syntetický hlas?

To se připravuje, to se chystá. Teď snad bude první hlas, pak by měl přibýt i ženský hlas. Pak snad i víc hlasů, aby se to dalo rozlišit, otázka, odpověď třeba, to se chystá a jsem na to hodně zvědavý, jak to bude vypadat v praxi, až to nastane, protože to bude další část, která se bude muset kontrolovat.

A co na to říkáte?

Je to určitě možnost. Ty lidi si budou chtít každopádně ten text taky poslechnout. Pokrok jde stále dál, už jsme na webu, už to dávno nejsou pouze tištěné noviny. Rozhodně to je cesta, jak jít nějakým směrem dál. Vlastně se tady čtenářům nabízí další možnost, jak se k tomu textu, nebo zprávě, dostat. Těch možností bude čím dál víc a myslím si, že už to není pouze o tištěné formě, i když prostě moje generace a starší ode mne, si ty noviny ráda ještě vezme do ruky. Teď už je to spíš o tom, že jde o více komplexní zprávu, která obsahuje text i video, co se dá podchytit. Třeba bude příště už i vonět...

Kdo by měl podle vás být odpovědný za obsah, který redaktor nebo editor vytvoří s pomocí umělé inteligence? Vy teda říkáte, že to nepoužíváte, ale kdyby vám například přišel nějaký článek od redaktora, který by AI použil...

Pokud se bavíme o zpravodajství, tak tam teď v tuhle chvíli podle mě ta umělá inteligence nemá co dělat. Samozřejmě můžeme to s ní zkusit, ale ta odpovědnost je vždycky na tom člověku, který to vydal. Který tu zprávu vypustil, ten má asi největší odpovědnost za to, co v ní stojí, jak vypadá. Ale to je obecně, myslím si, ať už je AI použitá nebo ne. Pomoc bude

spíš opravdu v té technické části, jak dostat tu zprávu ven. Nedovedu si to představit, že by úplně nahradila redaktora.

Říkáte tedy, že AI nepatří do zpravodajství, tak proč podle vás?

No právě proto, že vlastně já neznám ty algoritmy, nevím, jak by v tvorbě zprávy ten systém postupoval, podle jakých parametrů, nastavení a hodnot. A ve zpravodajství je důležité kdy, kde, co, jak a proč se něco událo a musí to být přesně. Nelze si tam prostě najednou něco vymyslet. Proto si myslím, že tam nepatří.

A to doporučování ano?

Doporučování ale není už v tom textu. Doporučování je většinou o tom postavení na homepage, nebo může nám to doporučit třeba související články, které se k tomu vážou. Ale ten text samotný, podle mě ne.

Takže to doporučování by podle Vás do zpravodajství, třeba na homepage iRozhlasu, patřilo?

Ano, de facto se dá použít. Já si můžu najít to doporučení, ale vždycky to konečné slovo má ten editor, který to má v tu chvíli na starosti.

Na začátku rozhovoru jste zmínil řešení ilustračních fotek. Ráda bych se k tomu ještě vrátila. O co tedy konkrétně šlo?

Podle našeho fotografa není úplně dobré používat „ilustračky“, protože se stává, že tam někdo někde najde například logo firmy a pak se nám dotčený ozve. Anebo někdo pozná na fotce dítě, a pokud je to focené v Česku ozve se, že si nepřeje použití tohoto záběru. To jsou věci, které se při používání ilustračních fotek stávají. Samozřejmě by se tomu asi předešlo, kdyby ty fotky vytvořila umělá inteligence. Ale zase na druhou stranu, protože jde o zpravodajství, tak je to jako kus falešné zprávy. Proto se tomu bráním. Teď jsme vlastně řešili, že to dělal Seznam a ten to dělal u komentářů, což je vlastně taky text, ale už to není úplně zpráva.

Mě ještě zaujalo, jak jste říkal, že jste to na tom iDNESu právě používal, tak jestli si dovedete představit, že by se to převedlo na ten iRozhlas...

Asi jo, to není nic složitého, nebo vlastně ono to tam už je, jen to tam nikdo nezaklikne v tom Chartbeatu, protože na té homepage vám pak straší taková různá čísílka, ale de facto tam vlastně tenhle zásah k dispozici máme. Není to teda přímo v tom systému, kde se edituje, ale je to podobné. A samozřejmě i když jsme veřejnoprávní, čtenost sledujeme.

Respondent Petr Jadrný, datum 25.10. 2023

Víte, co jsou to rekomenční algoritmy a jak fungují?

Ne.

Vzpomínáte si na řádek „Robot doporučuje na iRozhlasu“

Ano.

A co si o něm myslíte?

Přijde mi to jako zajímavá věc, která může pomoci, hlavně tedy jestli je to opravdu tak, že sleduje, co ten člověk čte a co má rád a podobně a doporučuje mu věci na základě toho, tak si myslím, že to může být dobrá věc.

Pracujete teď s automatickým doporučováním teď v rámci vaší editorské práce nebo ne?

Ne.

Kdybychom si teoreticky představili, že by vznikla stránka iRozhlasu jakoby druhá záložka, která by byla čtenáři ušitá na míru, podle toho, co už právě četl třeba, tak co byste na to řekli?

Mně by se to nelíbilo, protože mně to trochu připomíná ten problém, co se kdysi řešil ohledně těch nějakých krimi zpráv, a teď nevím, jestli to bylo na Nově nebo na Primě, kdy vlastně ti lidi říkali, že když něco těm lidem neukážete, tak oni nebudou vědět, že se to děje. Když bych dostával ke čtení jenom to, co vlastně čtu normálně, tak mě to hrozně ochudí o ty věci, a hlavně o informace, které anonymně... a teď to беру z pohledu jako čtenáře, ale kdybych to vzal z pozice běžného člověka, tak já sice chápu, že vás jako zpráva o novém rozpočtu nebo podobně zajímat nemusí, ani si jí nerozkliknete, ale už jenom to, že si člověk přečte ten titulku a ví, že se něco řeší, a dost často se v tom titulku i dozví tu nejdůležitější věc, co z toho potřebuje vědět. Nebo samozřejmě nejlepší je, když si to rozklikne, ale i tak se to jako dozví a má nějakou představu o tom, že něco takového existuje a že něco takového se řeší, a pak si to třeba může i rozkliknout. Ale když tu nabídku mít nebude a bude šitá jenom jemu na míru, tak mi přijde, že ho to ochuzuje a že to může být i dost nebezpečné.

Změnilo by to třeba nějak vaši editorskou práci?

To asi hodně, protože teď se vlastně snažíme nějakým způsobem vybírat ty články tak, aby prostě bylo pokryté všechno, aby prostě byla naplněna ta idea té veřejnoprávnosti, aby byly zastoupeny všechny možné názory a vybírat ty důležité věci, ale pokud by to vybíral automaticky počítač člověku na míru, tak... já jako nemám úplně představu, jak by to přesně jako konkrétně fungovalo. Ale na druhou stranu si myslím, že v tu chvíli vlastně já bych nemusel na tu homepage vybírat nic, já bych vlastně jenom ty články publikoval a pak by to vlastně všechno dělala ta umělá inteligence za mě. Takže myslím si, že tu mou práci by to změnilo hodně.

Kdo by za takový obsah nebo za takového doporučování měl být potom zodpovědný?

Velmi zajímavá otázka. Nemám nejmenší ponětí. Tak jako těžko z toho můžeme vinit ten počítač, nebo tu umělou inteligenci, teda, respektive v tomhle případě. Ale jako kdo přesně by za to měl být odpovědný... Tak vzhledem k tomu, že by to bylo na našich stránkách, tak asi my. Ale na druhou stranu jako neumím si to v praxi moc představit, jako že by to celé

bylo jenom podle těch lidí. No neumím si to představit, jak by to v tomhle přesně fungovalo, protože si myslím, že právníci by se docela zapotili.

A měl by čtenář vědět, že konzumuje obsah, který mu doporučuje algoritmus a jak byste mu to případně oznámil?

Asi by to měl vědět přesně z těch důvodů, o kterých jsem mluvil, že když prostě vím, že to něco je mi vybíráno na základě mého vlastního vkusu a na základě mých preferencí, tak je to něco jiného, než když je něco vybíráno obecně, vlastně pro populaci. Takže si myslím, že by to určitě mělo být dáno vědět. A myslím si, že ten způsob, jakým to máme udělané na iRozhlasu, že tam je vlastně přímo záložka a je to tam přímo jako vidět, tak to mi přijde jako dobré. Myslím, že je to dobrý způsob, jak to zobrazit.

A cítíte se případně na takovou práci, na takovouhle jako změnu v editorské práci připravený nebo ne?

Já jsem o tom ještě nepřemýšlel, ale nevím, asi bych to zvládnul. Nemám až takové problémy se něčemu přizpůsobovat, takže asi si umím představit, že bych s tím dokázal nějakým způsobem žít.

Říkám, otázka je, jak by to přesně fungovalo, protože si myslím, že jako zobrazovat těm lidem, jenom co je baví... třeba když to teď vezmu apriori na sebe jako na čtenáře, tak já třeba mám hrozně rád historii a obecně až takový ty bezpečnostní otázky a podobně. Ale jako myslím, že sám, i když třeba mi ekonomika absolutně nic neříká, tak třeba kdybych jako nevěděl nic o schvalování státního rozpočtu, nebo prostě o daňových balíčcích a podobně, tak by to taky nebylo asi úplně v pořádku. Říkám, záleželo by, jak by to bylo nastavené.

V komerčních médiích se díky rekomendačním algoritmům snaží třeba zvýšit právě personalizaci, tak má to používání rekomendačních algoritmů opodstatnění v Českém rozhlasu jako médiu veřejné služby?

Já myslím, že nemá. Tím, že by to vlastně popíralo tu myšlenku té veřejnoprávnosti a té nabídky pro všechny. Že by to sice bylo strukturované pro každého a vlastně my bychom stejně asi nemuseli pořád mít na webu články, které by jako splňovaly všechny ty požadavky. Ale myslím si, že vlastně my nepotřebujeme jakoby zvyšovat tu čtenost uměle. Samozřejmě čtenost je fajn věc, ale jako na druhou stranu, nemělo by to u veřejnoprávního média být jako jediné kritérium. Proto si myslím, že jako zajímavost nebo zlepšení je to pro člověka fajn, ale na druhou stranu si myslím, že ta homepage, ta hlavní část by pořád měla být ve stylu: „Tohle jsou věci, které byste si měli přečíst, protože si to editor nebo šéfredaktor nebo to médium myslí“. A pod tím by teprve mohlo být: „A tohle by vás mohlo zajímat.“ Takhle si to vlastně představuju já, ale nevím, jestli to není trochu zpátečnické.

A podle čeho se rozhodujete, když pracujete jako editor, tak podle čeho se rozhodujete třeba kde a co bude umístěno, jak dlouho to tam bude? Podle jakých hodnot třeba?

Toho je strašně moc... od stáří toho článku. Snažím se, aby nikdy nebyly dva články se stejným tématem nebo stejného formátu pod sebou. To znamená třeba příklad, teď jedeme dva onliny z ukrajinské války a z té dejme tomu izraelské války, i když tam je to trochu složitější nazývat válkou, ale rozumíme si. A já prostě nemám rád, když mám ty články „online dvojtečka“ dva pod sebou, takže prostě i tohle je jeden z důvodů podle kterých jako řadím ty věci na té homepage. Témata... pak vlastně záleží trochu i na tom autorství toho článku. Že pokud pokud je to článek z ČTK, tak pro mě, až tedy na samozřejmě nějaké

největší události a podobně, nemá takovou prioritu jako věci, které vzniknou v Českém rozhlasu. Ať už je to přímo na iRozhlase nebo ve vysílání. Nevím, těch kritérií je strašně moc, tak jako nevím, asi bych je všechny nevyjmenoval, ale to je asi tak to hlavní. Na té zpravodajské homepage je důležité, aby se ty články točily, aby to tam jak se říká, nezesmrádl, aby tam došlo k nějakým změnám, tak se to snažím točit. Ale jako těch věcí je opravdu hodně.

Takže to vy byste třeba zvládl líp než doporučovací algoritmus?

To si rozhodně nemyslím. Jako určitě bych... Já si spíš myslím, že to je jiná funkce. Že jako doporučovací algoritmus je něco, co mi má pomoci prostě užít si třeba jako den nebo prostě pomoci, abych dostal to, co mě baví, aby mi to vlastně i nějakým způsobem třeba jako zabilo můj čas, když to vezmou jako to... ale jakoby ta moje funkce jako editora je říct lidem: „Heďte tady, prostě je penzum, já nevím deseti článků, které jsou nejdůležitější z nějakého principu, který prostě máme nějakým způsobem nastavený, nebo protože si po těch několika letech, co to dělám, myslím, že jako mají nějaký smysl, tak proto tam jsou.“ Já si myslím, že tyhle dvě funkce jsou rozdílné. I z toho důvodu, jak jsem říkal, že kdybych vlastně dostával jenom ty věci, které jsou pro mě, tak jako jsem ochuzený o spoustu věcí, které by mi to nenabídlo.

No teď to není jako doporučovací algoritmus, ale vím, že pracujete s takovým tím tagováním nebo že vybíráte vlastně manuálně, že jo, související články. Je to tak?

Ano.

Tak když byste si představil, že by vám vždycky, když editujete nebo dáváte na web nějaký článek, takže by vám k tomu bylo nabídnuto automaticky třeba, nevím dalších deset doporučených článků, tak by vám to nějak pomohlo v tom výběru, nebo naopak?

Myslíte jako že když by vybíral ty články, které tomu čtenáři nabídnu a aby v tom článku jako pokračovali dál, tak kdyby je nabízel počítač a bylo by to tak, že jich nabídne deset a já si čtyři vyberu, nebo jak je to myšlený?

Já jako nemám tu přesnou představu, ale spíš jde o ten princip toho, jestli by vám to ušetřilo nějaký čas nebo práci.

Ono by hodně záleželo na tom, jak to bylo nastavené a jak by to bylo dokonalé, protože my takhle podobným způsobem vybíráme klíčová slova. My tomu říkáme našeptávač. Je to vlastně jako tlačítko, na které kliknete a ono vám to doporučí nějaká klíčová slova.

Jako tam si myslím, že ani zdaleka nikdy nevyberu všechny. Vyberu jich jako půlku a pak jich tam třeba ještě jednou tolik dost často dopisuju ručně. Takže jako mně osobně to třeba docela pomáhá a mám to docela rád. Takže umím si představit, že kdyby to bylo dobře vymyšlené, dobře nastavené a nabídlo mi to třeba 10 článků a já bych si z toho prostě 3 nebo 4 vybral. Nebo my tam obvykle dáváme 3, takže 3 vybral, tak si myslím, že by to byla fajn věc. To bych jako ocenil a přišlo by mi to docela fajn a opět záleží na tom, jak by to bylo nastavené.

A máte třeba nějaký nápad, jak by to mohlo být nastavené tak, aby to pro vás bylo co nejefektivnější?

Tak ono by to muselo být... jako takhle, přemýšlím, jak by to muselo fungovat i jako technicky. Ono by to muselo asi vybírat nějak ty články podle nějakých klíčových slov, aby jich bylo co nejvíc podobných a zároveň v čase. A pak je otázka, do jaké míry by to mohlo

fungovat. Nebo možná by ta umělá inteligence mohla číst i ty články a přiřadit je k tomu. Nevím, neumím si to představit. Tohle už je bohužel trošku nad moje technické schopnosti.

Tak z toho pohledu toho editora, co by vám pomohlo bez ohledu na to, jestli to technicky lze nebo nelze provést.

No mě by pomohl ten předvýběr. Já teď ty články vybírám tak, že vyberu jedno, maximálně dvě klíčová slova a podle toho vybírám ty související články a vlastně když to řeknu blbě a jako línej člověk, tak vlastně mě to stojí vlastně čas, kdy já musím tam ty dvě klíčová slova zadat, vybrat je z toho seznamu, a pak je to teprve vyhledává. Takže když bych já jenom kliknul na tlačítko a ono mi to rovnou vyhodilo těch deset článků a já bych prostě z těch... ono by to víceméně tenhle krok udělalo za mě, nebo třeba by to bylo ještě chytřejší a dokázalo by to ten článek přečíst, ten co, co zrovna já dodělávám nebo edituji a dokázalo by to přečíst ty staré články a vybrat z nich těch 10, které by se tam mohly hodit a já bych si z toho 3 vybral, tak by to prostě pomohlo.

Napadají vás u toho nějaké výhody... teda, to jste už jste vlastně řekl, ale nějaké třeba etické výzvy, které by se s tím mohly pojít.

S tím výběrem těch článků? Etický výzvy asi moc ne... jakože umím si představit. Ono, jak jsem tady už naznačoval u toho vybírání těch klíčových slov, tak tam se prostě stává, že zaprvé ten systém nám občas vymýšlí klíčová slova, která nám nevyhovují třeba tou svou formou, protože on je nastavený tak, nebo mně to přijde, že je nastavený tak, aby byl co nejpřesnější. Takže nám třeba ruská jména nabízí i s tím otčenstvím. Myslím, že se tak říká tomu prostřednímu jménu. A to mi prostě nepoužíváme. Prostě málokdo použije termín Vladimír Vladimírovič Putin a tak dál.

Takže já si neumím moc představit, že by tam byl nějaký etický problém, protože je, stejně jako pokud by to ta umělá inteligence jenom vybírala a nabízela to tomu člověku mně jako editorovi, nebo i mně jako redaktorovi, to už by bylo celkem jedno a já bych si z toho jenom vybíral, tak jsem za to stejně pořád odpovědný já. Takže tam by bylo jediné co, ledaže by to někdo nastavil, nebo by se ta umělá inteligence rozhodla, že z jeho etických jako parametrů se rozhodne nějaké články nebo témata nenabízet a nebylo by to schválené tím člověkem. To je asi jediné, co mě napadá.

Ted' si nejsem jistá, jestli jsem se na to už neptala...

Tak se zeptejte klidně znovu.

Tak jestli by čtenář měl vědět, když má nějaký ten obsah doporučován algoritmicky, tak jestli by v tom případě měl vědět, konzumuje takový obsah.

Ano. Myslím si, že ano.

A jak byste to třeba označil nebo jak byste mu to dal vědět?

To už jsme se bavili, že já bych to dělal úplně stejně a že mi to řešení, jaké máme na iRozhlasu, přijde fajn. Že tam prostě máme tu lištičku, tu záložku s tím „zprávy, které jste nečetli“. Jinak si umím představit taky, že by se to dalo dělat nějakým jiným oddělením, že by to mohlo mít podbarvení. My máme bílé, tak by to mohlo mít nějaké jiné. Nevím, těch variant by bylo asi hodně, ale mělo by to být dost výrazné na to, aby to bylo vidět.

Měl byste zájem o školení nebo nějaký manuál od projektových manažerů ohledně doporučovacích algoritmů?

Určitě, já mám tahle školení a nové věci a nástroje, se kterými si člověk může tak jako pohrát, tak mám rád. I kdyby se ukázalo, že to k ničemu praktickému úplně není, tak se na to rád podívám. Takových nástrojů jsem viděl spoustu, a i když některé z nich byly spíš jako taková zajímavost, která se v praxi nebo ostrém provozu úplně neosvědčila, tak si myslím, že to, že to nevyužiju teď hned, tak neznamená, že to nevyužiju třeba v budoucnu. Nebránil bych se tomu a za něco podobného bych byl rád. Tohle není úplně o školení, ale naše projektová manažerka tam s námi sedí denně nebo Honza Cibulka... někdy je to školení, ale spíš je to tak, že prostě přijdou nebo pošlou mail s tím, že „tady máme novou věc a něco, co by vám možná mohlo pomoci, koukněte, tady je návod nebo si to zkuste nebo se za námi stavte“.