

## UMĚLÁ INTELIGENCE A HUDEBNÍ TRADICE

*Matěj Kratochvíl*

„Vytvoř píseň ve stylu cajun na text novinové zprávy o hasičích zachraňujících bobra uvízlého v plotě.“ Stručná slovní instrukce spojená se zkopírováním textu z náhodně zvolené, lehce bizarní zprávy ze zpravodajského serveru, necelá minuta čekání a webové rozhraní nabízí k poslechu přibližně dvouminutovou píseň odpovídajícího stylu s českým zpěvem poměrně správně frázujícím. Software sám rozdělil text na sloku a refrén, odhadl, kde má vložit pauzu, aby byla zachována logika plynutí vět, nenechal se zmást ani anglickým slovíčkem, jež se do výchozí zprávy vloudilo. Intuitivní ovládání a rychlá cesta k výsledku stojí za popularitou služeb jako SUNO či Udio, které nabízejí generování hudby na základě slovních instrukcí uživatelů.<sup>1</sup> Díky jejich masovému rozšíření během let 2023 a 2024 se pojem umělá inteligence stal součástí každodenních diskuzí mezi laickými uživateli a jimi vytvořené kompozice začaly být široce sdílené. Podobný byl tento nárůst zájmu u programů generujících obrázky, jejichž produkce začala v posledních dvou letech zaplňovat veřejný prostor. I u hudby lze předpokládat, že s rozvojem generativních programů bude stále větší podíl hudby kolem nás vytvářen pomocí algoritmů. Dokazují to spekulace o využívání těchto systémů streamovací službou Spotify. Z pohledu etnomuzikologie pak tento trend přináší dvě zásadní otázky. Přichází s těmito technologiemi něco nového do způsobu, jak přemýšlíme o hudbě a s ní spojených tradicích? A jak nové tyto technologie ve skutečnosti jsou? V následujícím textu se pokusím nastínit možné odpovědi na první otázku a skrze historickou exkurzi se zastavím u té druhé, abych se dostal ke známému zjištění, že pod sluncem není nic tak docela nového.

1. Webová stránka suno.com byla spuštěna v prosinci 2023, udio.com v dubnu 2024.

Zmíněné generativní modely jsou schopné pracovat s nej-různějšími druhy instrukcí, nejčastější je však kombinace popisu požadovaného instrumentálního obsazení s hudebně-žánrovými charakteristikami. Systém generující hudbu vychází z takzvaných velkých jazykových modelů, tedy souborů dat opatřených indexy. Ve vztahu k hudbě to tedy může být vzorek určité nahrávky označený jako „cajun“. Pokud bude ono slovo v mém zadání, systém se daný vzorek pokusí napodobit.<sup>2</sup> V praxi se ovšem ukazuje, že jakkoliv se generativní modely prezentují jako schopné pokrýt veškerou hudbu, materiál, na němž byly natrénovány, je výrazně selektivní. Snaha o napodobení některých tradic tak nevede k úspěchu. Například označení „lidová hudba / folk music“ u systému SUNO spolehlivě vede k výsledkům vycházejícím ze severoamerické tradice country či bluegrassu. Jakékoliv upřesňování nedokáže vyvolat asociace například s evropskými tradicemi.

Žánrová označení představují specifický příklad hudební tradice. To, že hudbu s určitými znaky označujeme jako cajun, polku či verbuňk, je výsledkem společenského konsenzu v daném místě a čase. Klíčovou otázkou je zde fungování paměti a mechanismy jejího přenosu, tedy to, co lze považovat za základ tradice. Jak popisuje americký folklorista Henry Glassie, „*tradice je vytvářením budoucnosti na základě minulosti. Tradice coby průběžný proces situovaný v nicotě přítomnosti a spojující zmizelé s neznámým je zastavena, rozparcelována a kodifikována mysliteli, kteří se zaměří na ten či onen aspekt podle svých potřeb či zájmů a zanechají nám směs zjevně si protirečících, nicméně přesvědčivých definic*“ (Glassie 1995: 395).<sup>3</sup> Takto se kodifikují postupy jak hudbu tvořit, ale také jak jí označovat a kategorizovat.

Tradice, její plynutí, přerušování a navazování, je tématem mnoha vědeckých prací. Pro potřeby této úvahy je vhodná studie etnochoreologa Andriye Nahachewského (2001), v níž se její autor zamýšlí nad vývojem lidového tance. Navazuje na starší

2. Podrobnější popis fungování generativních systémů např. <<https://www.bmat.com/what-really-is-ai-music/>>.

3. Všechny překlady z angličtiny autor.

práce muzikologa Felixe Hoerburgera (1968) a rozvíjí různé variace vývoje a ožívování tradic. Podstatné je, že Nahachewsky reflektuje, jak do vývoje tradic zasahují např. interpreti či sběratelé. Zjednodušeně můžeme tradici popisovat jako proud informací, do něhož se zapojují různí aktéři. V hypotetické první existenci folkloru se tento proud odvíjí mezi jednotlivými generacemi, jež si předávají hudební repertoár a další související vědomosti. Od 19. století pak do hry stále výrazněji vstupují další hráči, sběratelé, vydavatelé jejich sbírek, někdy též cenzoři a další. Každý, kdo se podílí na plynutí tradice, má vlastní motivace a je ovlivňován danými společenskými podmínkami. Na to odkazuje i biolog a hudební folklorista Vladimír Úlehla, když ve své knize *Živá píseň* hledá paralely mezi zkoumáním lidové hudby a přírodními vědami, konkrétně ekologií: „*A ani v přímém okolí nositelově neudržoval se vzhled písňového porostu stále stejný. Vlastní podnebí písňové – celková dispozice zpěvákova – rozhodovalo o tom, zda se přízeň údobí vegetačního rozdílela rovnoměrně mezi všechny členy zpěvákova repertoáru, či zda se některé časem rozbujely, časem zas byly přehlšovány jinými.*“ (Úlehla 1949: 221) Proměny vnějších podmínek se pak podle Úlehly vždy propisují do hudby samotné: „*Obracím se nyní k nitru strážnické písňe a chci hledat v něm doklady o účasti podmínek přírodních, nositelových i společenských na utváření toho, co písni především zdá se propůjčovat osobitosti, tj. básnického obsahu a nápěvného toku.*“ (Úlehla 1949: 277)

Datové balíky zmiňovaných velkých jazykových modelů, na nichž stojí generativní umělá inteligence, mohou při prvním pohledu budit dojem pouhé modernější verze písňových sbírek. Jde rovněž o soubor hudebních informací, nějakým způsobem utříděných. Povaha generativních systémů je ovšem oddělená od onoho Úlehlova „podnebí písňového“, je zafixována teoreticky na věky, pokud bude existovat počítač schopný data rozluštit. Hudba je zde uchována v dokonale zmrazené podobě. Zatímco přístup živých lidí k tradici zaznamenané do písňových sbírek se v čase výrazně proměňuje, data jednou vložená do počítače nám vydají obraz nedotčený plynutím času. Na rozdíl od písňových archivů v sobě umělá inteligence spojuje archivaci hudebních dat a jejich

interpretaci. Jako kdybychom se do edice Karla Jaromíra Erbena či Františka Bartoše nemohli sami podívat, ale museli bychom si zde otiskované záznamy vyslechnout v podání zpěváka, jenž jako jediný má k původním záznamům přístup a jako jediný má možnost je interpretovat.

### Československé počátky strojového generování

Kořeny zájmu o využití počítačů v hudební tvorbě lze hledat v období po druhé světové válce, kdy rozvoj technologií zasáhl široké pole lidských činností. Od počátku šlo o doménu hudby vážné, jejíž představitelé hledali nové cesty k opuštění dědictví romantické estetiky a nástroje pro experimenty. Za klíčové dílo lze považovat *Illiad Suite for String Quartet*, kterou Lejaren Hiller a Leonard Issacson nechali v roce 1957 vzniknout na počítači ILLIAC I. Podobných pokusů následně vznikla celá řada, většina z nich s poměrně malým dopadem na posluchačstvo. Lejaren Hiller se následně stal autorem řady textů dokumentujících postupný vývoj a proměny role počítačů v hudbě. Dařilo se mu mapovat dění po celém světě a díky tomu zaznamenal i pokusy v socialistických zemích včetně Československa. Dva z nich jsou dnes spíše zapomenuté, v souvislosti s úvahou o minulosti a budoucnosti umělé inteligence v hudbě ovšem zasluhují pozornost. Tím spíše, že právě lidová hudba se v nich stala hlavním materiálem.

Antonín Sychra (1918–1969) bývá v textech o dějinách oboru zmiňován především pro svou publikaci *Hudba a slovo v lidové písni* (1948). Tento estetik a jeden z poválečných propagátorů estetiky socialistického realismu se v 60. letech stal propagátorem využití kybernetiky, tehdy relativně nového oboru v hudební analýze. Ve sborníku *Nové cesty hudby* vydaném roku 1964 najdeme jeho článek *Hudba a kybernetika*, v němž demonstruje principy počítačové analýzy na moravských lidových písních vybraných z edicí Jana Poláčka, Františka Bartoše a Leoše Janáčka. Ukazuje přitom důkladnou znalost dobového stavu bádání, včetně přehledu západní literatury. Sychra neměl k dispozici počítač, proto svou „kybernetickou hru“ realizoval s pomocí kartotéčních lístků, na něž si vypsával jednotlivé melodické a rytmické prvky písní.


Každý prvek opsal v tolika kopiích, kolik odpovídalo jeho procentuálnímu zastoupení v Sychrou analyzovaném vzorku. Následným losováním z těchto kartiček Sychra zkoušel vytvářet nové melodie, z nichž několik je též ve studii otištěno. Badatel je označuje jako „danaje“<sup>4</sup> a následně rozebírá tvůrčí limity takového přístupu (Sychra 1964: 241). Sychrova studie zapadá do kontextu dobové fascinace novými technologiemi, které poskytovaly vítaný únik od dřívějších ideologických diskuzí o správné socialistické hudbě.

Identita tvůrce skutečně prvního pokusu o strojově generované melodie vycházející z dědictví lidové hudby zatím zůstává zčásti nejasná, stejně jako okolnosti, za nichž k němu došlo. Zdeněk Fencel byl zaměstnancem Výzkumného výpočtového střediska národního podniku Kancelářské stroje. V roce 1966 vydal v tehdy čerstvě vzniklém časopisu *Kybernetika* článek nazvaný *Komponující algoritmus a obsah informace*. Velice stručně jej otevírá sdělením, že jde o snahu „*dospět analyticko-syntetickou metodou k hudební kompozici stochastickým procesem a využít zkoumané látky k analýze z hlediska teorie informace*“ (Fencel 1966: 243).

Základem se Fencelovi stalo osmdesát sedm písní vybraných ze sbírky Karla Jaromíra Erbena, přičemž rozhodujícím hlediskem pro selekci byla malá dvoudílná forma s reprízou prvního závětí. Tento vzorek byl dále analyzován po stránce formy, harmonie, rytmu a melodie a tato analýza se stala zdrojem informací o pravděpodobnosti výskytů jednotlivých kombinací hudebních prvků. Důležitým pojmem byl pro Fencela Markovův řetězec, tedy pravděpodobnostní proces popisující přechod z jednoho stavu do druhého, v tomto případě od jednoho tónu k dalšímu. Zjednodušeně řečeno Fencel zanalyzoval necelých devět desítek písní z jedné historické sbírky a na základě této analýzy vytvořil algoritmus schopný vygenerovat neomezený počet dalších písní odpovídající stejným principům. Ve své studii otiskl čtyři z těchto nových melodií, k jejichž vytvoření použil v té době výjimečný výdobytek západní technologie, britský počítač National Elliott 803 B.

4. Označení *danaj* používá Antonín Sychra ve svém textu poněkud vágně, víceméně ve smyslu párového točivého tance ze slováckého Dolňácka.

OPUS 01100100110010111000

1   
T D T T T T D T T T T T T D T

OPUS 00100010010011111111

2   
D T D T D T D T D T D T D T D T

OPUS 10100101010001010010

3   
T D T T T D D T D T T T T D D T

OPUS 01010011000010001010

4   
T S D T T S D T T D T D T S D T

*Písňe vygenerované prostřednictvím algoritmu vytvořeného Zdeňkem Fenclem.*

*Zdroj: Fencel 1966: 236*

Národní podnik Kancelářské stroje byl jedním z míst, kde i během socialismu bylo možné udržovat kontakt s vědeckým a technologickým vývojem kapitalistických zemí. Hlavní motivací bylo zlepšování fungování státních podniků a tím i národního hospodářství, prostor ovšem byl i pro na první pohled ekonomicky málo přínosné pokusy. O motivaci pro volbu tématu Fencel nic neuvádí a v době vzniku tohoto textu nebyly dostupné žádné informace o jeho případném hudebním vzdělání. Studie upadala do zapomnění, zřejmě i kvůli svému názvu, který nijak neanoncuje souvislost s lidovou hudbou. Zároveň se ale informace o ní dostala k výše zmíněnému americkému skladateli a badateli v oblasti hudebního využití počítačů Lejarenu Hillerovi. Ten ji (a také pokusy Antonína Sychry) stručně zmiňuje v historickém přehledu vydaném roku 1979, vedle podobných pokusů ze Sovětského svazu a Maďarska. Podle Hillerova přehledu se zdá, že přinejmenším v té době bylo využití výpočetních metod k práci s lidovou hudbou doménou socialistických zemí.

Sychra i Fencl byli ve svých experimentech motivováni spíše analytickým zájmem, nicméně principy, které při tom využili, jsou velice blízké těm, s nimiž pracují dnešní generativní modely. Sychra si navíc ve svém textu s myšlenkou skutečně kreativního využití kybernetiky při tvorbě hudby na několika místech pohrává, byť podle něj má modelování tvůrčího procesu především význam badatelský: „*Přitom není nesnadné využít zkušeností a pravidel, vytěžených ryze empiricky nebo statisticky z hudební tradice. Je-li nadto k dispozici analogový vstup, tj. může-li stroj zpracovávat ne binární nebo jiný matematický kód, nýbrž přímo akustický materiál, není nesnadné prověřit různé zvukové realisace; to je podstata moderních zvukových syntetisátorů.*“ (Sychra 1964: 244) Tato vize je již velmi blízká aktuálnímu stavu, kdy hudební generativní software dokáže kombinovat práci s algoritmy a konkrétními zvukovými záznamy. Vývoj v české hudební vědě a etnomuzikologii se nicméně ubíral jiným směrem a počítače našly uplatnění především v oblasti analýzy. Tam s jejich pomocí došlo k významnému vhledu do struktury českých lidových písní (Tyllner – Vejvoda 2019) či do budování databáze lidových písní a tanců.<sup>5</sup>

### **Viditelní a skrytí dveřníci**

Proč se obracet k článkům zapomenutým po desetiletí, když přemýšlíme o nejnovějších trendech? Jejich obsah i okolnosti vzniku totiž upozorňují na důležitou vlastnost tradice, a to jakékoliv. Co do ní patří a co ne, je vždy výsledkem komplikovaných procesů, do nichž zasahují různě motivovaní aktéři. Dnes, stejně jako v 60. letech je pro generování nových melodií zásadní, na jakém materiálu je systém vytrénován, z čeho odvozuje pravidla pro chápání daných hudebních struktur. A kdo jsou lidé či instituce rozhodující o tom, co se do datových balíčků dostane a co nikoliv, jaké hudební informaci se dostane jakého označení. Dobrou cestu k lepšímu pochopení těchto otázek nabízí teorie *gatekeepingu*.

5. Viz portál <lidoveprameny.cz>.

První použití tohoto pojmu bývá spojováno se sociálním psychologem Kurtem Lewinem (1943), do podoby propracovaného systému jej rozvedla informační teoretička Karine Barzilai-Nahon. Podle její definice *gatekeeping* označuje „*způsob kontroly informací prostřednictvím jejich výběru, přidávání, zadržování, zobrazování, usměrňování, formování, manipulace, opakování, načasování, lokalizace, integrace, ignorování a vymazávání*“ (Barzilai-Nahon 2008: 1499).

Ke *gatekeepingu* dochází ve společnosti na mnoha úrovních, v jakékoliv situaci, kde proudí informace. Často má toto rozhodování neformální podoby a mnohdy zůstává nereflktováno. Nevyhýbá se ani sběru a výzkumu lidových písní, ať už jde o sběratele individuální, či institucionální. V případě jednotlivých sběratelů hraje roli osobní vkus, představy o tom, co je „ta správná“ lidová píseň, jak můžeme pozorovat v textech Leoše Janáčka a dalších. I zde může hrát roli politika, jak popsala ve svých studiích Lucie Uhlíková (2012, 2015). Politika pak vystupuje výrazněji do popředí, díváme-li se na institucionálně organizované sběratelské akce, jako byl například výzkum hornických komunit na Kladensku v 50. letech 20. století (Kratochvíl 2022).

Jak konkrétní *gatekeeperi* (či česky dveřníci) – ať individuální, či institucionální – ovlivnili, co a jak má být z lidové hudební kultury zachováno? To lze zpětně zkoumat na základě přímých i nepřímých historických pramenů, jsou-li dostupné. V případě Kladenska tak lze porovnat záznamy písní dochované v archivu s tím, co bylo oficiálně publikováno a z toho si udělat obrázek, jak dveřníci ovlivňovali zpracovávaný materiál. Kdo je však v roli dveřníků v případě generativních modelů umělé inteligence a jak lze jejich práci zkoumat? Systémy využívající generativní umělou inteligenci, které mají potenciálně největší dopad na globální hudební život, jsou většinou soukromé společnosti, jež nemají zájem zveřejňovat detaily svých technologií. Jedním z důvodů jsou i právní ohledy. Existují podezření, že jako zdrojová data, na nichž se hudební generátory učí a která jsou tedy základem jejich obchodního modelu, je využívána hudba chráněná autorskými právy. Majitelé generativních programů ovšem nikomu neplatili a odmítají ukázat,



s jakými daty pracují. Zatímco tedy historii a případné zkusobní sběratelů i oživovatelů lidové hudební a taneční kultury můžeme do nějaké míry kriticky zkoumat, u těch, kdo hudbu významně ovlivňují nyní, jsme odkázáni na spekulace.

Další pohled na etické problémy spojené s využitím umělé inteligence ve spojení s prameny tradiční hudby přinesl experiment uskutečněný ve Velké Británii. Zatímco Zdeněk Fencel ani Antonín Sychra neměli ambice výsledky svých experimentů šířit a zařazovat je do kontextu skutečných lidových písní a nechat je veřejně hrát, badatelé Bob L. Sturm and Oded Ben-Tal se o to pokusili s materiálem založeným na anglických a irských melodiích. V roce 2018 vydali album nahrané živými hudebníky, ovšem s melodiemi většinou vygenerovanými počítačem. Skutečný původ použité hudby ovšem nejdříve nepřiznali, protože jejich cílem bylo právě zkoumání posluchačských reakcí na algoritmičticky generovanou hudbu. Toto rozhodnutí i celá představa počítačově vytvořených melodií napodobující tvarosloví tradiční hudby vyvolalo v britské komunitě výrazně odmítavou reakci. Nabízí se otázka, co by se stalo, pokud by takovéto „lidové“ melodie vytvořené umělou inteligencí ve větší míře získaly úspěch mezi hudebníky a staly se součástí živé kultury. Vzhledem k množství digitalizačních projektů realizovaných po celém světě v institucích spravujících sbírky lidových písní nemusí jít o nerealizovatelnou představu.

Umělá inteligence je termín, pod který se schovává velké množství systémů a aktivit. V souvislosti s hudbou se zkoumá její možné uplatnění k analýze gregoriánského chorálu, k automatizovanému přepisu notového zápisu či generování zvuků k počítačovým hrám. Popisované systémy pro snadné generování hudby podle slovního zadání nejsou vědecky ani kreativně nejzajímavější aplikací. Zdá se však, že jde o aplikaci se zásadním dopadem na hudbu, která nás obklopuje v každodenním životě a která ovlivňuje naše uvažování o ní. Právě z toho důvodu bude zásadní, aby etnomuzikologové a badatelé ze spřízněných oblastí tomuto vývoji věnovali kritickou pozornost.

\* Text vznikl s podporou na dlouhodobý koncepční rozvoj výzkumné organizace RVO: 68378076 (Etnologický ústav AV ČR, v. v. i.).

## Literatura:

- BARZILAI-NAHON, Karine 2008: Towards a Theory of Network Gatekeeping: A Framework for Exploring Information Control. *Journal of the American Society for Information Science and Technology* 59 (9): 1493–1512.
- FENCL, Zdeněk 1966: Komponující algoritmus a obsah informace. *Kybernetika* 2 (3): 243–258.
- GLASSIE, Henry 1995: Tradition. *The Journal of American Folklore* 108 (430): 395–412.
- HILLER, Lejaren 1970: Music Composed with Computers – A Historical Survey. In: LINCOLN, Harry B. (ed.): *The Computer and Music*. Ithaca: Cornell University Press. 42–96.:
- HOERBURGER, Felix 1968: Once Again: On the Concept of “Folk Dance”. *Journal of the International Folk Music Council* 20: 30–32.
- KRATOCHVÍL, Matěj 2022: *Vy havíři umouněný, co vy z toho máte... Výzkum hornických písní na Kladensku v letech 1953–1959 v kontextu dobové vědy a politiky*. Etnologický ústav: Praha.
- NAHACHEWSKY, Andriy 2001: Once Again: On the Concept of “Second Existence Folk Dance”. *Yearbook for Traditional Music* 33: 17–28.
- SYCHRA, Antonín 1964: Hudba a kybernetika. In: *Nové cesty hudby. Sborník studií o novodobých skladebných směrech a vědeckých názorech na hudbu*. Praha: Státní hudební vydavatelství. 234–267.
- TYLLNER, Lubomír – VEJVODA, Zdeněk 2019: *Česká lidová píseň. Historie, analýza, typologie*. Praha: Bärenreiter.
- UHLÍKOVÁ, Lucie 2012: Právě perly písňové: edice českých lidových písní a rakouská cenzura. In: PŘIBYLOVÁ, Irena – UHLÍKOVÁ, Lucie (eds): *Od folkloru k world music: Hudba a bariéry*. Náměšť nad Oslavou: Městské kulturní středisko. 18–27.
- UHLÍKOVÁ, Lucie 2015: Duch a povaha národa v písní. Idealizovaný obraz lidové písně v tištěných sbírkách první poloviny 19. století. In: WÖGERBAUER, Michael – PÍŠA, Petr – ŠÁMAL, Petr – JANÁČEK, Pavel a kol.: *V obecném zájmu. Cenzura a sociální regulace literatury v moderní české kultuře. 1749–2014*. Praha: Academia – Ústav pro českou literaturu AV ČR, v. v. i., 2015. 333–344.
- ÚLEHLA, Vladimír 1949: *Živá píseň*. Praha: F. Borový.

## Elektronické zdroje:

- STURM, Bob – BEN-TAL, Oded 2018: „Let’s Have Another Gan Ainm: An experimental album of Irish traditional music and computer-generated tunes.“ *KTH* [online] [cit. 15. 10. 2024]. Dostupné z: <<https://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A1248565&dswid=6358>>.

## *Summary*

### **Artificial Intelligence and Musical Tradition**

Generative AI offers a new form of music creation where users can input instructions and receive songs in a desired style and instrumentation, based on text-based instructions. This phenomenon has led to the rapid proliferation of services such as Suno and Udio that generate music based on algorithms and patterns. However, these technologies often fail to emulate specific musical traditions, raising questions about their influence on musical culture. The text shows how the understanding of traditions is transformed by the rendering of large language models used by artificial intelligence. It shows the historical attempts at machine generation of music in Czechoslovakia in the 1960s by Antonín Sychra and Zdeněk Fencel, who used folk music as their material. Their methods are close to contemporary generative systems that combine music analysis and algorithms. It shows how important were and are the “gatekeepers” who decide what musical elements are preserved in the tradition. In current AI models, these gatekeepers are private companies that hide the details of their training material, calling into question their transparency and ethics.

**Key words:** artificial intelligence, folk music, transformations of folk tradition, technology