

AIで学生は創造的になれる（音楽）

植田 康孝*

要 旨

言語モデルとオープンソースの組み合わせが爆速でコンピューターの進化を生み、音楽分野も凄いスピードで進化している。プロの作曲家が試行錯誤しながら苦労して作った候補作2曲から選ぶより、AIが1時間で作った1,000曲の中から最高の素材を探す方が、結果的に良質の作品が出来る。AIがランダムに作曲するため、曲が出来るペースより、人間が聴く方が時間を要するため、永遠に追い付けない。出来上がった曲が不満であっても、即座に代替曲を提示することが可能である。「夏の出会いが出て来る音楽をアイドルソング調で作って」とプロンプトに入力すると、きちんと感情を揺さぶる音楽を創る。しかも、それが数秒で且つ無料で出来てしまう。言語モデルは年に1回くらいアップデートされるため、2035年には人間の作成能力を超える。AIの動画生成は作曲を上回るペースで進化を遂げている。動画のBGMは見過ごされがちであるが、需要が多い分野である。TikTokが大ヒットした理由は、BGMが予め用意されているからである。絵画では、人工知能（AI）の生成画像をベースにした作品が審査員に気づかれないまま有名な賞を獲ったが、音楽のグラミー賞でもAI作成曲は対象外とすることが発表された。特別ルールを設けなければ、人間の製作を超えてしまう証左である。人間が作った音楽とAI生成曲で明確な区別が付くのであれば、このようなルールは必要としない。音楽を創る時間とコストは猛烈に下がり、質はどんどん向上している。2023年末時点では、「Suno」の評判が良い。

生成型AIを用いた実習のポイントは「創作の楽しさ」が学生の学習に対するモチベーションにもたらす効果である。音楽を作って発表すること自体がまず楽しい。そして全員集合型で「苦しみながら頑張った」よりも、楽しみつつ個人としてオリジナルの創作物が出来る実感を持ち易い。生成型AIをきっかけに多くの学生に創作の楽しみを知ってもらい、「鑑賞」するだけに過ぎなかった受動接触から、「製作（参加）」する能動接触に変えられる。結果、メディアやエンターテインメントを学習するモチベーションが高まれば良い。個人が苦手とした部分をAIがカバーしてくれるため、音楽の知識や楽器スキルがなくても、作曲家・作詞家、ミュージシャンになれるようになった。生成AIの用途の多くは画像や動画など視覚的なものである。様々な生成AIを実習してもらう中で、聴覚を刺激する「音楽生成AI」は特に学生たちが興味を持って取り組んで好評であった。ボカロPやVチューバー人気に伴い、自らが作詞・作曲・演奏・歌唱しなくても音楽の情報発信をするハードルは極端に低くなっている。音楽生成AIは、合成音声を使ってゼロから音楽を作成できる。音声空間、音声の生成、変換、編集、ダビング、翻訳などの機能は、私たちが作る多くのものの基礎となるメディア要素になる。生成AIは音楽教育を他人が作ったモノについて蕋蓄を垂れるだけの評論家思想からクリエイティブ思想に変えてくれた。「創る」ことは楽しい。そして最先端の「知」に触れる機会でもある。

キーワード：クリエイティブの民主化 特殊技能の大衆化 坂本龍一 小室哲哉 サカナクション・山口一郎
YOSHIKI ビートルズ新曲 AIチューバー ボカロP YOASOBI「アイドル」 Suno

1. 音楽生成 AI の登場

生成AIは、文章生成と画像生成で目を見張る

2023年11月30日受付

* 江戸川大学 マス・コミュニケーション学科教授
理学博士（国際情報通信学）

進化を遂げている。そして現在、注目されているのが、動画生成、3次元モデル生成、そして音楽生成である。音楽生成に関しては、グーグル、メタ、マイクロソフトが既に研究開発を進め成果を公表した。音楽生成AIは音楽や音声を自動生成するAI技術である。本分野で良く用いられるアルゴリズムには、WaveNet、敵対的生成ネット

ワーク、LSTM（長短期記憶）があり、リアルな人間の声を模倣したり、新しい音楽トラックを生成したり、様々な音響効果を作り出すことが可能である。また音楽分野の生成 AI スタートアップも増えており注目が高まる。生成 AI の登場により誰もが音楽製作できるようになり、音楽製作者が消費者へ直接作品を届けられる時代になれば、テレビ局や音楽レーベル、音楽配信プラットフォームの価値は薄れる（図 1）。ネット空間での主導権が個人や新興企業に再び引き戻される。ソニーグループは音楽を「鑑賞」して楽しむエンターテインメントから「創造」するエンターテインメントへシフトする時代認識を踏まえ、「クリエイターに寄り添う会社を目指す」と掲げた。AI 時代には商品やサービス、アイデアを生む「クリエイター」をどこまで自陣に引き寄せられるかが企業戦略上で重要と判断した。

音楽生成 AI には 2 つの方向性がある。一つは、作詞作曲できる人を更に上手く作曲作詞作曲できるようにすること、楽器を演奏できる人を更に上手く弾けるようにすることである。他一つは、作詞作曲を学ぶ、楽器演奏を学ぶ初心者向けマーケットにおいて、ユーザーが手を伸ばし易くして興味を持たせ、初心者でも音楽を生み出せるようにすることである。音楽生成 AI は、初心者でも自ら楽曲を作り出し、練習して演奏できる点である（図 1）。音楽生成 AI はテキストを入力したりメロディーの一部を入力したりすれば、楽曲を作り出してくれる。ユーザーが 1 フレーズ弾けば、続きを 1 フレーズ作ってくれる。

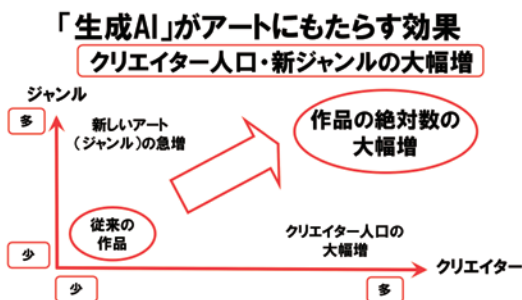


図 1 音楽生成 AI による特殊技能の大衆化 (I)

出所：筆者（植田）が作成

AI の登場で変わるのは、コンテンツ量が爆発的に拡大することである（図 1）。ライブやコンサートなど高額チケットを購入してくれる人や音楽配信に月額課金を支払ってくれる会員だけに音楽を聴いてもらえば良いという考え方は危うい。一般の関心や機会が減れば中長期的な音楽ファンを失いかねない。生成 AI は完全なコピペではなく、何か入れたら新しいモノが出て来るツールである。画像生成 AI に対してイラストレーター（絵師）が反発しているのに対し、音楽生成 AI は技術的に未熟であり、人間を脅かす水準まで至っていないため、アーティストからの反発は少ない。音楽分野では目下、新しい表現、収入源になるというポジティブな捉え方が多い。音楽には元々ヒップホップ、マッシュアップ的な文化があり、受け入れられ易い土壌が存在する。CM 音楽や短期的に納品を求められるアーティスト以外は、生成 AI を新しいシンセサイザー、新しいピアノとして捉え、生成 AI を使って新しいアートをクリエイトする文化が起きている。レコードが登場した時、生演奏じゃなきゃ駄目という反発が起きた。しかし、Spotify などの音楽配信、YouTube などの動画で音楽が聴けるようになった現在、ライブ、コンサートはむしろ価値が上がっている。テイラー・スウィフトのワールドツアーが各地でインフラを起こすなど影響は大きい。二分化が加速している。

2. 音楽生成 AI をめぐる動向

エンターテインメントやメディア業界はグローバルに飛び抜けた力がないと競争に勝てない。世界の音楽業界は、CD やダウンロードの所有権を手放し、サブスクやストリーミングモデルへと大きく舵を切った。CD 売上枚数よりもストリーミングによる消費量に力点が置かれるようになった。30 秒以上ストリーミングされなければ収益化されないため、アーティストは「曲は短く、聞かせどころは最初に」というように曲作りの方法論すら変化を余儀なくされ、ベテラン教員による音楽教育が時代遅れとなった。日本市場だけが

CD が売れる恵まれた市場として残存したのは、CD を押しグッツ化したためであり、音楽メディアとしての用途ではない。そのような日本でも普及する音楽配信では、「ボカロP」や「Vチューバー」などの影響により、ヒット曲の常識が変わった。Ado（ウタ）、ヨルシカなど顔やプロフィールを公開せずに音楽活動するアーティストが増え、ヒットソング上位100曲の1割を占める。ボカロPのコンポーザーであるAyaseが作る楽曲をボーカルのikuraが歌唱するYOASOBIが2023年4月12日にリリースした「アイドル」は2023年6月10日、米ビルボードの北米発以外グローバルチャートにおいて日本語で歌われた楽曲として初の1位を獲得した。ストリーミング累計再生回数も史上最速で2億回を突破した。2022年10月からアニメ「チェンソーマン」の主題歌「KICK BACK」を配信して日本語の歌詞として全米レコード協会から初めて「ゴールド」（米国内での記録的ヒットを示す）に認定された**米津玄師**も当初、ボーカロイドで楽曲を作るボカロPとしてほぼ顔を出さず活動し人気につなげた。顔を出さずに「誰でも自由に楽曲を発信できる時代」の先駆者と言える。

人工知能を使うことで若者や学生も音楽の創作に参加し易くなり、音楽の裾野が急速に広がった（図1）。もはや音楽を受動的な視座から蘊蓄を垂れる社会学や人文学、経営学のアプローチは「いつの時代の教育？」と評価されるようになった。受講学生から次のヒット曲を生み出すアーティストが生まれるかもしれない。実際、メジャーレーベルに所属せず活動するインディペンデント（独立）系アーティストの躍進が続く。レコード会社の力を借りずとも、個人で発信した音楽がヒットチャートを賑わす。現在の音楽シーンにおいては、オーディション選考やメジャーデビューを経ずとも成功を収める例が当たり前になった。音楽生成AIによりハードルが下がり、インディペンデントアーティストの活躍の裾野が広がったことで、多様なジャンルから次世代のスターが登場する可能性が加速している。

音楽業界における生成AIに対する姿勢は、著

作権や倫理的問題を指摘し、その利用を制限すべきという否定的な立場と生成AI活用を促進する肯定的な立場の2つに大別される。生成AIを創作の現場が避けることは難しい。実際のところ、サカナクションの山口氏、故坂本龍一氏、小室哲哉氏、YOSHIKI氏など新規技術に挑戦的なアーティストは多く、生成AIを早速、創作現場に適用し始めた。「新宝島」などのヒット曲を持つ「**サカナクション**」ボーカルの**山口一郎**（2020）氏は音楽生成AIについて「僕が幼少期から書き続けて来た文章をAIに学習させてみようというプロジェクトを一時期やっていました。僕はAIが作り出す音楽というものに対してすごく肯定的です」「テクノロジーって様々な感情のミックスを行って来るものだと思うんです。人間が作ったものと同じ感動を生み出すかは分からないんですけど、新しい感動を生み出してくれるような気がします」と言う⁽¹⁾。

音楽家の**小室哲哉氏**は、2022年3月から理化学研究所の客員主管研究員に就任した。小室氏は「AIで作曲活動を効率化できる」と述べた。早くからAIの可能性に着目していたと言う小室氏は、自身の音楽制作への技術導入を目指し、AIによる音楽制作支援の在り方を探求すると言う。理化学研究所の**浜中雅敏博士**（2023）は「AIが人の仕事を奪うのでは」という警戒感に対し「単調作業をAIに任せれば、作曲家は芸術性を高めるための創作活動に集中できる」と言う。浜中博士はAIで音楽の構造を分析したり作曲を支援したりする「音楽AI」の第一人者で2023年3月には前述の小室哲哉氏を客員主管研究員として招き、AIを使った作曲支援システムを共同開発すると発表した。三重県桑名市は2023年8月21日、2026年4月開校予定の小中一貫校「市立多度学園」の**校歌**をAIで作詞作曲すると発表した。地元の子供や保護者たちから校歌に盛り込みたいキーワードを募集し、理化学研究所が開発したAIに学習させる**AI超校歌**～AIがつくるみんなの校歌」に、歌詞に盛り込みたいキーワードや曲調を学習させれば、校歌の原曲が完成できる。

2023年3月28日に亡くなった**坂本龍一氏**は千

葉工業大学長・伊藤穰一氏との対談で「僕も人工知能の利用については基本的にポジティブなので、人工知能が作曲することは構わないし、出来上がった音楽が面白かったら評価はするでしょう。でも、その音楽がどういうプロセスで、どのようなプログラムによって作られたかの方が大切だと思う。モーツァルトの曲を人工知能に全部覚えさせて、統計的に似たような曲を作るのは簡単にできるじゃない。それでは面白くないよね。少なくともその音楽の美の本質が何なのかというのをアルゴリズム化しないと。人工知能は、人間が思いつかない、人間には絶対不可能な量の計算が得意なので、そこから何が生まれるかに期待したい。出来上がった音楽を僕らが聴いて面白いと思えば、それは恩恵だと思う」と言っていた。

YOSHIKIは2023年2月23日、スタンフォード大学で公演し、米グーグルが音楽を自動で生成できるAIを開発していることに触れ、音楽というクリエイティブ産業に大きな可能性を与え始めていると話した。生成AIは「私たち自身を破壊するものではなく、お互いを支え合い、愛するために使わなくてはいけない」と述べた。会場に集まったスタンフォード生にどんな言葉を贈りたいかと聞かれたYOSHIKIは、「be your own rock star（自分自身のロックスターであれ）」と呼び掛けた。

ワーナー・ミュージックが音楽生成AIスタートアップ「LifeScore」に投資を行うなど積極的な姿勢を見せる。「LifeScore」は、Siriの共同開発者であるトム・グルーバー氏が設立したとして注目されるスタートアップである。2023年5月、韓国の歌手「イ・ヒョン」氏は、韓国語や日本語など計6か国語で新曲を同時発表した。母語の韓国語以外の5か国語は人工知能を使い、自然な発音に修正した。「MIDNATT」という歌手名で発表した新曲「Masquerade」は韓国語、日本語の他、英語、中国語、スペイン語、ベトナム語でそれぞれ流暢に歌う「イ・ヒョン」が6か国語で歌うMVは下記で見られる⁽²⁾。

<https://www.youtube.com/watch?v=l19AEmRvh-Y>

「ユーミン」の愛称で知られるシンガー・ソングライター、松任谷由実が2022年「デビュー50周年」を迎え、楽曲「Call me back」では、AIで再現したデビュー当時の荒井由実の歌声と出演した。MVでは、ユーミンがはるか未来で歌う「AI 荒井由実」に会いに行く物語が紡がれる。「私が将来いなくなっても、AIがきっと曲を作ったり、パフォーマンスをしたりしている。私のアルゴリズムがAIによって成長させられ、人体を離れた所で色々なことをやって行く」と50周年記念ベスト盤の特典映像で語った。音楽プロデューサーである松任谷正隆は2023年9月自身のラジオ番組（FM東京）で「コンピューターがもう作曲できる時代なんです。もっと細かくAIが学習していけばこの先、人が作りがちな曲も当然作曲できます。『Call me back』が当たり前の世界になります」と伝えた。

優秀なアーティストは今でも大人による管理を好まず、生産的な切磋琢磨を音楽業界に与えてくれる。テクノロジーを駆使すれば、どのような業界にも革命を起こせることを示し、ストライキに踏み切った映画などテクノロジーに否定的な業界とのスタンスの違いを際立たせた。ハーバード大学のマイケル・サンデル教授（2023）は「最近、ポール・マッカートニー氏がAIを使って古いデモテープからジョン・レノン氏の声を抽出し、ビートルズの新曲「ナウ・アンド・ゼン」を完成させた。AIが発達すれば、彼らが作曲してさえない『ビートルの新曲』を彼らの声やスタイルで作ることが出来るだろう」と音楽業界の姿勢を高く評価した⁽³⁾。「ビートルズ」の元メンバー、ポール・マッカートニーはジョン・レノンが残した音源を基にビートルズの曲を完成させ、2023年11月2日に27年ぶりの新曲「ナウ・アンド・ゼン」を世界同時発売し2023年11月10日には全英シングルチャート1位を飾った。全米シングルチャートでも7位を記録した。11月17日に発売された日本盤は11月27日付オリコン週間ランキングで6位を獲得した。「レット・イット・ビー」（1970年5月）以来53年7か月ぶりのトップテン入りだった（図2）。音楽業界には、ビートルズ

から更なる音楽が生まれるというアイデアを歓迎する人たちが少なくなかった。NHK 紅白歌合戦に登場した「AI 美空ひばり」には「死者への冒瀆」を反発する声が上がった⁽⁴⁾一方で、欧米では AI を使いビートルズの新曲を出すことにそこまでの反発はなかった。死生観やデジタル表現の受け止め方の違いなどが背景にある。

AI 技術を使って、雑音の入った録音テープからレノンの歌声だけを取り出し、加工・編曲した。ポール・マッカートニーがベースやピアノ、リンゴ・スターがドラムを重ねて録音し、ジョージ・ハリスン（2001 年死去）が 1995 年に録音したギターの音を使い、完成させた。AI 音声処理ソフトは不快なノイズの軽減や除去だけでなく、様々な音声の修復が簡単な操作で行われるようになった。例えば、唾液によってピチャピチャと鳴ってしまった唇や舌による雑音も、クリック一つで綺麗に消すことが出来る。そのような音声修復

AI を使ってデモ音源からジョン・レノンの歌声だけを取り出し、新たな演奏を加えて製作された。ビートルズは活動を通して常に当時の最新技術を駆使したレコーディングを手掛けて来た。この先の音楽とテクノロジーとの関係を示唆するものになったと言える⁽⁵⁾。

2023 年 6 月 16 日、米音楽界最高の栄誉とされる「グラミー賞」を主催するレコーディング・アカデミーは、人間の作者を含まず AI だけで作った楽曲は授賞の対象外とする新ルールを設けた。2024 年の第 66 回グラミー賞に向けた新ルールで、「授賞の対象となるのは人間の創作者のみとする。人間の作者を含まない作品はどの部門でも対象外」と定めた。AI 技術で生成した素材を使った音楽でも、人が主体となって創作した場合は、特定の部門で授賞の対象となる。ソングライティング（作詞や作曲）系の部門では、大半が人間によって作られた楽曲である必要があり、パフォーマンス部門も人間のパフォーマーが対象になった。

3. 音楽生成 AI を使った実習

生成 AI に対しイラストレーター（絵師）や俳優・脚本家からクレームが上がったのに対し、音楽生成 AI に対してアーティストから激しいクレームは起こらなかった。画像生成 AI が誰もがダイレクトに見える高画質画像を生み出したのに対して、音楽生成 AI は技術的にそこまで到達しておらず反発も少なかった。音楽生成 AI を用いた実習のポイントは「創作の楽しさ」が学生の学習に対するモチベーションがもたらす効果である。音楽を作って先生や学生の前で発表して聞いてもらえること自体がまず楽しい。そして全員参加型で「苦しみながら頑張る」従前の学習スタイルよりも、楽しみつつ主体的に作ったオリジナルの創作物が出る実感を持てる。生成 AI をきっかけに、多くの音楽を受動的に楽しむに過ぎなかった学生に創作の楽しみを知ってもらった。音楽を聴くのは好きだけど、作詞や作曲、演奏は難しそうだから止めておこうという学生は多い。高校時代に吹奏楽部であったが、大学に入ると音楽自体を

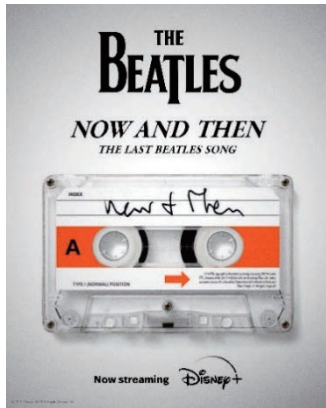


図 2 「ナウ・アンド・ゼン」

<https://www.youtube.com/watch?v=Opxhh9Oh3rg>

やらなくなる学生も少なくない。それらの学生にも生成 AI は大きく下げてくれ、音楽を創る楽しさを与えてくれる。

音楽の要素として、メロディー、ハーモニー、リズム、テンポ、拍子、楽曲の編成、曲の展開、エフェクトなどがあるが、音楽生成 AI ではこれらの要素を自由に選択できるが、音楽の基礎がないと、難しかったようである。結果として「鑑賞」するだけに過ぎなかった受動接触から、「製作（参加）」する能動接触に変えることで、メディアやエンターテインメントを学習するモチベーションが高まれば良い。生成された音楽は、まだまだ人間のアーティストと比較すると見劣りするため、人間の作品を代替するモノと言うより BGM が主であり、人間の創作活動や創造的生活をサポートする存在に近い。チャット GPT などテキスト（文字）生成 AI を用いれば、歌詞やギターコードを生成してくれるが、本稿では作曲 AI を対象とした。音楽を好きな人にも「歌詞を聴く」派と「歌詞を聴かない」派がいる。「歌詞を聴かない」とはメッセージや意味を能動的に受け取る意識がないというニュアンスである⁽⁶⁾。

3.1. SOUNDRAW (サウンドロー)

「SOUNDRAW」は、作りたいイメージや尺感を入れると AI が自動的に音楽を作ってくれる自動作曲サービスである。音楽の知識がないユーザーでも楽曲を簡単に作れる。ジャンルやムードを選ぶだけで、AI によって生成されたフレーズを組み合わせて楽曲が自動生成され、動画に合う構成・尺の BGM⁽⁷⁾を簡単に作ることが出来る。「旅行」や「ビジネス」といったテーマ、「ハッピー」や「ドラマティック」といったムード、「ヒップホップ」や「ロック」といったジャンルを選択して生成ボタンを押すと、10 曲以上の楽曲が6秒程度で生成される。各楽曲はテンポ、楽器、曲の構成、長さなど簡単な操作で変更できる。生成した楽曲のイントロやサビを手軽にアレンジ出来る機能が受け 2023 年 5 月時点で 3 割が日本、残り 7 割は北米を中心とする海外が占める。

AI が音楽を作るサービスは数多くリリースさ

れているが、本 AI の特徴は、

- (1) テーマを決めて行くと、楽曲を AI が提案
- (2) テンポ、楽器、キーを変更できる
- (3) 楽曲の中でどこを盛り上げるか変更可

などが出来る点である。AI が共通の特徴を抽出し、特徴を持つ大量のバリエーションをクラウド上に置き、組み合わせて楽曲として提供する。5,000 万通り以上の楽曲を生成できる。

SOUNDRAW (サウンドロー) 公式サイト

名称	URL
SOUNDRAW (サウンドロー)	https://soundraw.io/ja/

日本のスタートアップ企業による AI ツールである。曲の長さやテンポ、ジャンル、ムードを指定するだけで 10 曲以上を自動生成する。曲中の一部でメロディーやドラムの強調具合を変更するなど曲と映像のマッチングを調整できる。ユーザーがボーカルや別の楽器の演奏を加えるなどオリジナル要素を加えれば、楽曲販売するなど BGM 以外の用途でも利用できる。2023 年 8 月 24 日のテレビ東京「WBS」で特集されたほど注目された、2023 年 8 月 25 日には大手レーベル（ソニー）より生成オリジナル曲が配信された。

学生の感想①

「暗く、負の感情が渦巻くような曲編成にしたかったため、壮大な印象を与えるシネマティックを選び、テンポは重く感じる事が出来るように遅めにした。曲の主となる部分にピアノは入れず、ハーブを用いた。ハーブは細い音が出る印象があり、強い音の楽器を多く用いるのではなく、ドラムとのバランスを取るためピアノを外した。テンポを 80 から 90 に変更するだけで、曲が持つ重さが全く異なるものに出来た」(図 3)

学生の感想②

設定を「ダイナミック」「ロック」にして AI に生成させ、ゲーム BGM のようなループ系の音楽をイメージして調整した。面白いと思ったのは、楽器の音量を調整できる点である。自分が好きな楽器を目立つように音量を大きくしたり、逆に必

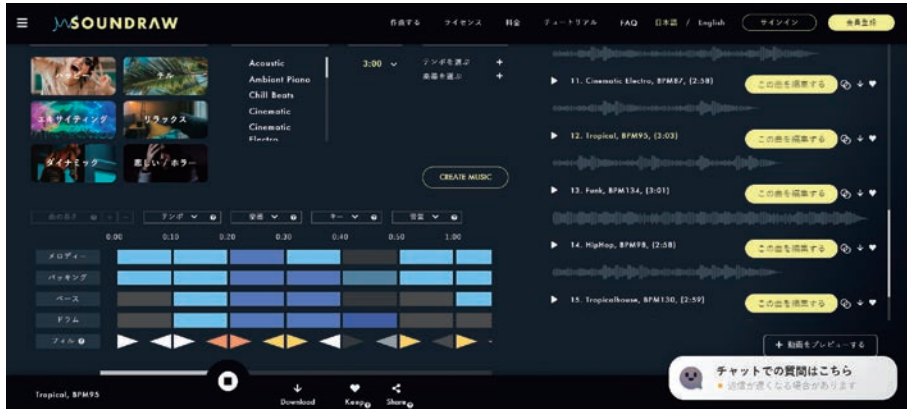


図3 SOUNDRAW（サウンドロー）（学生①）

要ない楽器だけ小さくしたりというカスタマイズ性が気に入った。私は軽音部でベースを弾いているので、ベースが目立つよう調整した。難しい調整や音楽の構成を考える必要がなく、誰でも作曲を簡単に楽しめると感じた。

学生の感想③

「私は SOUNDRAW を使って曲を作りました。一番好みの曲は、ギターがメインになっているもので、前向きになれる仕上がりです。テーマにハッピーを選んで、長さは1分、テンポを速いものにしました。AIが作った15の候補曲を聞きましたが、どれも満足の行く出来でした。ここまで来ると誰かに歌って欲しいと思いました。作曲家になったみたいで、気分が非常に良かったです。楽曲作りは音楽の知識を持っていなければ出来ないと、勝手に決めつけていました。今後、本格的にやりたいと思いました。」

学生の感想④

「今回私は soundrow を使って曲を作った。

- ムード チル
- ジャンル スタイル HIPHOP
- 長さ 1分
- テンポ 普通

を選択して AI に曲を作ってもらった。AI が作った曲が15曲出て来た。自分好みの曲に仕上げていくために

- テンポ 120
- 楽器 ベース シンセリード

• キー K2

に設定した。楽譜も読めない自分に曲を作ることが出来るのか不安だったが、ジャンルを選ぶだけで自分好みの曲を作ることが出来た。今まで作曲は才能ある人や曲を作る技術を持っている人のみができることが出来ると思っていたが、私のような素人でもしっかりと曲を作ることが出来た。私はボーカロイドのメロディーがとても好きなので、シンセリードやキーを全体的に高くして曲を作った。」

3.2. SOUND（サウンド）

名称	URL
曲の自動生成	https://tomari.org/main/java/audioapi/audio_sakkyoku.html

学生には、サウンドロー以外の音楽生成 AI をいくつか紹介した。全ての AI を試すというよりも、自分の好みに合う、使い易い AI ツールを選択するよう指導した。

学生の感想⑤

「入力するだけで簡単に曲を作ることが出来ることに感動した（図4）。作曲は、音を入力したり、自分で楽器を弾いたりして出来るイメージがあり、自分には到底できない印象だった。しかし、自動生成により自分で簡単に曲を作ることが出来た。アップテンポな曲で疾走感を感じられるよう作曲した。シンセサイザーで疾走感を感じる

自動生成設定

生成拍数: 、拍数/分(BPM/MM):

音階 (☐ 個別指定 ☒ 和音コード指定)

小節指定 (☒ 全小節 ☐ 四小節毎)

分類	コード	生成	分類	コード	生成
和音	C	4	音符	全音符	10
	Dm	1		付点二分	6
	Em	4		二分	
	F	3		付点四分	
	G			四分	4
	Am	7		付点八分	
	Bm ⁻⁵			八分	2
休止			休止符	全休符	
				付点二分	
				二分	
				付点四分	
				四分	1
				付点八分	
				八分	
			十六分		

図4 曲の自動生成「SOUND」(学生⑤)

ことが出来た。ピアノで聞くと、シンセサイザーよりも柔らかくなり、違った印象を受けた。」

3.3. 「オルフェウス」

名称	URL
「オルフェス」	http://www.orpheus-music.org/

嵯峨山茂樹・東京大学名誉教授のチームが開発した「オルフェウス」は、歌詞を入力すると、AIが歌詞の音の高低を解析し、ふさわしい曲を付け機械音で歌い上げてくれる。歌声や楽器、曲の速さを選択すると、AIが作曲してくれる。最後に「作曲」のボタンを押すと、10秒で曲が流れる。

(1) ユーザー登録する

管理者が読んで審査するため、通常、数分～数日後にパスワードがメールで届く。ユーザー登録（無料）に記入し、数時間～数日後にパスワードが送られて来たら、ログインする。登録したプロ

フィールが登録者一覧に表示される。パスワードが送付されて来るタイミングが見通せないため、ゼミナールや演習実習では、本AIを唯一の選択とせず、他の音楽生成AIも紹介して、タイムロスが生じないように配慮した。

(2) 自動作曲する

日本語歌詞から自動で作曲し、伴奏付きの合成音声で歌う。その後に二重唱も作れる。

(3) 自動作詞する

数個のキーワードから歌詞を自動で作り出してくれる。

学生の感想⑥

「作曲した感想は、DTMや打ち込みをPCで仕上げる人に最適だと感じました。植田先生も話していたように、ボカロ曲はエレクトリックな機械音を扱うので、そのような音楽作りをしている人に向いていると感じました。メリットは、伴奏などベースとなる音源もあるので、コードが分からなくても、自動生成してもらってから修正を利かせることが出来る点です。メロディーやピッチも自由に変えられます。自分もギターをやっている曲を作りたいと考えたことが何度もありました。これを機会に作曲に挑戦してみます。」(図5)



図5 自動作曲「オルフェウス」(学生⑥)

学生の感想⑦

「初音ミクが歌いそうな曲を作ろうと思い、歌詞を厨二病臭いものにした。歌詞は、『マカロン、イチゴのショートケーキ、とりあえず君にぶち込んで 胃もたれパーティー グンナイパーティー 明日なんて誰も来ない〜』として、ボカロの歌詞

に出て来そうな漢字の歌詞を書いた。童話に出て来そうな曲調の曲が出来上がった。声もボーカロイド風であった。何回もリピートしていると、だんだん癖になるメロディーになった。受講する以前は、ゼミの先生に『生成 AI を使ってはいけない』と言われていたので、AI は人間を脅かす道具であり頼ってはいけないイメージを持っていたが、AI を使って色々なものを作成してみて、自分の発想力では補えない部分を AI にやってもらうことは、能力を伸ばす上で必要であると思った。作曲できたので、プライベートでも新しい曲をどんどん作って行こうと思う。植田先生と AI に背中を押してもらえた。」(図 6)

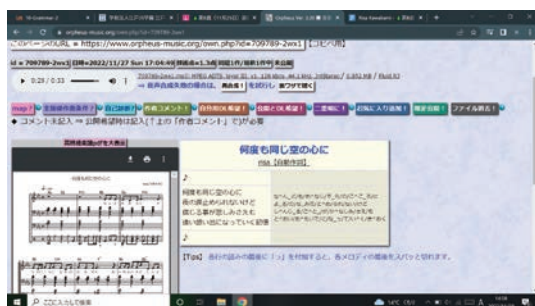


図 6 自動作曲「オルフェウス」(学生⑦)

学生の感想⑧

「人生で初めて作詞作曲をしました。貴重な経験でした。もっと学びたいと思いました。音楽に携わる仕事がしたいと思っていたのですが、これから興味ある音楽を製作サイドから学ぶようにします。家の中が寒くて暖房を付けるかここ 3 日くらい悩んでいる気持ちを書いたら、凄くユニークな曲になりました。音楽を好きだと言う学生は



図 7 自動作曲「オルフェウス」(学生⑧)

たくさんいますが、聴いているだけでは仕事につながらないことを気付かせてくれた実習でした」(図 7)。

学生の感想⑨



図 8 自動作曲「オルフェウス」(学生⑨)

「作詞作曲したのは『狂役令嬢』です。最近、異世界の恋愛ものにハマり、ヒロインを意地悪する悪役令嬢に憧れ作りました。しかし、『悪役令嬢』ではなく『狂役令嬢』です。若者で 10 人に 1 人はいる地雷系の女の子を歌詞に盛り込もうと作詞しました。三角関係で王子と仲の良いヒロイン令嬢に嫉妬する感情を盛り込んでいます。サブの楽器にオルゴールを掛けゾクゾクする怖さを演出したり、テンポを落として不吉にしたり、狂っている感じを演出しました。機械音声ゆえ歌詞と音が調和し仕上げられました。」(図 8)

学生の感想⑩

「櫻坂 46 の曲のタイトルで日本語の単語をキーワードとして入力し、あとは自動で作詞作曲しました。完全に自動で作曲しましたが、音楽知識がある人が手を加えたら凄く良い曲になると思いま



図 9 自動作曲「オルフェウス」(学生⑩)

した。こんなに簡単に出来るなんて面白いと感じました。」(図9)

3.4. 「SongR」

Webブラウザから英語詞を入力するだけで、AIが作曲して歌詞入りの楽曲を自動生成する米国のサービス「SongR」が話題になった。

<https://app.songr.ai/>

類似AIは従来もあったが、ポップスやピアノロックなどから曲調が選べることや、歌声が自然なことを特徴とする。アメリカのスタートアップ「RIFFIT」が開発したAIである。無料でユーザー登録も不要で作曲させることが出来る。トップ画面に、ポップス、ピアノロックの他、ヒップホップ、カフェという4つの曲ジャンルが表示されるため、ジャンルを選択する。作詞もAIにしてもらいたい場合、真ん中の項目に歌詞や単語を入力して矢印ボタンをクリックする。既に歌詞を用意している場合、下にある「I have my own lyrics」をクリックして歌詞を入力する。RENDERをクリックすれば準備完了する。英語で750文字以内のテキストを入力した上で、「エミリー」(女声)か「ケビン」(男声)を選んで30秒待つと、入力した歌詞と曲調、歌声で歌を自動生成する。生成歌はダウンロードして公開できるため、国内外で多くのユーザーがX(旧Twitter)で曲を公開している。

3.5. CREEVO(クリーボ)

名称	URL
「CREEVO (クリーボ)」	https://creevo-music.com/ https://www.kagua.biz/markettool/20230225a1.html

音楽生成AIは日本でも以前から研究され、使い易いものとして「CREEVO」(クリーボ)がある。京都大学が2021年5月3日にリリースした。ユーザーの評価を基にして新しいAIモデルを生み続けることを特徴とする。CREEVO(クリーボ)は、京都大学の中村栄太特定助教の構想を元に開発された自動作曲AIシステムである。名古屋工業大学の歌声合成技術「Sinsy」を活用する。

日本語詞をベースにメロディー・伴奏・ボーカル(歌唱合成「SongR」)・動画まで自動生成してくれる。日本語詞を入力するだけで、自動的に作曲できる。構成やリズム、和声(コード)進行などカスタマイズ可能である。また、テンポやキーメロディー、サザン風とかミスチル風、あいみょん風などに変えたり出来る。無料で著作権フリーであるため、SNSで使える利点がある。例えば、日本語で作った歌詞を英語にしてSongRに歌わせた後、CREEVOに歌わせた。日本語の歌詞はCREEVOの方が自然に聞こえるため、日本語歌詞で作曲したい人には、3.3項で紹介した2007年に東京大学が開発した「Orpheus」よりも「CREEVO」がおすすめ出来る。

4. ボカロ曲・歌声生成AI

4.1. ボカロ曲

2023年時点、「YOASOBI」「Ado」「米津玄師」など、日本のアーティストで海外において評価される大半がボカロ系になっている。また、日本の音楽シーンでは、アニメ、ゲームのキャラクターやVチューバーが登場する音楽ライブが目立つようになっている。テレビ東京はAIとVチューバーを組み合わせた「AIチューバー」を育成する事業に進出した。Pictoria(ピクトリア)は「AIチューバー」の事務所「AI CAST」も設立した。「魔法少女アイリーン」などのAIチューバーの配信もスタートした。Vチューバーに関わる音声合成やバーチャルライブについて各種の実習を行なったが、本稿では紙幅の関係上、音声合成の導入部分のみを紹介し、AIチューバーの詳細は別稿に任せたい。

4.2. VTuberのボイスチェンジ

YouTubeやSNSなどインターネット上のコンテンツから登場したアーティストがチャートを賑わすことは珍しいことではない。日本においてその先駆けとなったのが「VOCALOID(ボカロ)」と「ニコニコ動画」である。「ボカロ」とは、2003年にヤマハが開発した「VOCALOID(ボー

カロイド)」の愛称である。ボカロ文化の象徴であり、実質的にスタート地点ともなったのは、「初音ミク」の誕生である。その人気と共にボカロの国内市場価値は拡大した。音楽ジャンルの一つとして、あるいはカルチャーの一つとして、ボカロは無視できない存在となった。実際、ここ数年で登場したアーティストの中にも VOCALOID クリエイター「ボカロ P」出身は数多く存在しており、バンドを組む、オーディションを受けてシンガーになる、という音楽を生業とするための入り口の一つとして効果的に機能している。むしろ、PC 環境とソフト、一定の DTM スキルさえあれば、幅広い層にクリエイターへの道を切り開くことが出来る。音楽シーンを見渡せば、ボカロに端を発した才能が数多く存在する。もともとコンピュータソフトであったこともあり、ボカロ曲は生成 AI と非常に親和性が高い。学生にはボカロ曲生成の実習にも取り組んでもらった。一部を 4.3 項に紹介する。

4.3. 「VoiSona」(ボイソナ)

名称	URL
「VoiSona」(ボイソナ)	https://voisona.com/

無料で提供する音声創作ツール「VoiSona」(ボイソナ)は、ユーザー登録するだけで誰でも無料で使うことが出来る。テクノスピーチは 2022 年 9 月 1 日、AI 技術で人間の歌声をリアルに再現する音声創作ソフト「VoiSona」を無償リリースした。VoiSona は、最新の AI 技術で誰でも手軽に人間のような歌声を再現できる歌声合成ツールである。デフォルトで中性的な声が魅力のボイスライブラリとして「知声 (Chis-A)」が同梱され、「VoiSona」と「知声」を用いて出力した音声波形データを無料で利用できる。音量やピッチ、ビブラートの振幅・周期、声質、ハスキー度をコントロール出来る。2023 年 8 月 9 日にはテクノスピーチから、「知声」のバージョン 2.0 を提供するメールを受け取った。以前のバージョン 1.X 系より高品質な音声出力が可能となっており、「知声 2.0」も無料で利用できる案内であった。AI

ツールがアップデートされ度に、案内が来るので「時代遅れ」になることはない。学生の感想は、下記の通りである。

学生の感想⑪

「ティックトックで流行っている 8LOOM の『Melody』という 15 秒の曲と、ティックトック編集機能にあるテキスト読上げ機能の AI 音声、ユーチューバーの話す声の 3 つを録音して、違う声の系統からボイスチェンジを試しました。エフェクトでコーラスやエコーなど様々な種類から音声を自由に変えることが出来、動画編集している気分になりました。エコーのプリセットから、洞窟と選択したら洞窟の響き渡るような声に変換され面白かったです。AI 音声は、機械声から加工声に変わり、不思議な感覚でした。アバターに歌を歌わせボーカロイドで有名アーティストとのコラボが出来ると感じました。」(図 10)

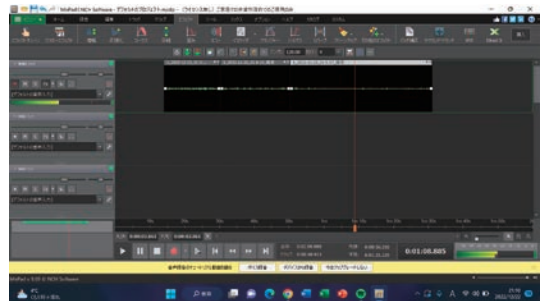


図 10 「VoiSona」(ボイソナ) (学生⑪)

学生の感想⑫

「音階を見ながら『君が代』を作ったが、歌とマッチする音階を探すのに苦労した。本当の歌のようにテンポや声の伸びを作ることが難しかった。しかし楽譜を見ながら曲を作るのはとても楽しかったので、他の曲にも挑戦したい。」(図 11)

面倒臭いと思った作業は、一音ずつ発声する音の入力だった。音程が整ったら、また始めに戻ってダブルクリックして一音ずつ入力することに時間を取られた。1 つずつ入力して全部発音が揃って再生できると感動を得た。一気に入力出来ないか探すと、「歌詞まとめ入力 (F8)」の方法を見つけた。試すと、まとめ入力はオススメ出来る機能であった。音痴や楽譜など関係なく誰でも音楽を

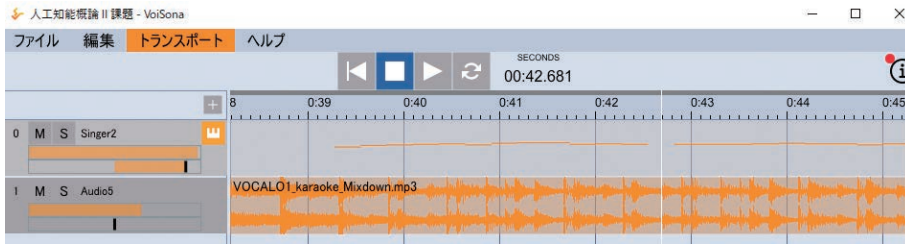


図 11 「VoiSona」(ボイソナ) (学生⑫)

作れる時代になったと実感した（図 11）。

学生の感想⑬

以下、2つのことに挑戦した。

①自己紹介を作る（個人的評価：初級レベル★）

- 自分の声に近い音声合成は，自分の音域を理解していないと難しい。
- 読ませるセリフが事前に出てくると，30分程で仕上げられた。
- 助動詞「は」は，「ha」と読まれてしまうため，「わ」と入力する必要があった。同様に「へ」も「え」と入力する必要があった。
- イントネーションの調節が合わないとボカロロイドより片言感が気になる。
- 英語では，一音にこだわらず単語ごとで音を合わせると，ネイティブっぽくなった。
- 「Hi（ひ）」と入力すると，「ハイ」と読まれてしまった。

②曲を作る（個人的評価：初級レベル★★★）

- 音程，音域，ビブラート，音量など細かい作業が密集しているため，思っていたよりも簡単ではなかった。
- 操作自体が細かい音域の組み立てをする感覚だったので，パズルのように楽しいと思った。
- 音痴で上手く音程を掴めないからこそ，初期段階は「音声入力で波形を作ってくれたら楽だろうな」と感じた。自分なりのアレンジを加えるともっと楽しくなると思った。
- 効率を求め PC よりもタブレットで操作したくなり，作業中にピンチアウトやタッチペンを使いたくなる衝動があった（図 12）。



図 12 「VoiSona」(ボイソナ) (学生⑬)

学生の感想⑭

人間の歌い方をリアルに再現できる AI を使った感想は，誰でもボカロ P になれる時代になった事だ。私の姉は初音ミクが好きで姉の車で出掛ける時には，ボカロ P が作成した曲ばかり聞いている。まさか自分がボカロ P 側になるとは思ってなかった。音声創作 AI を使ったボカロ創作は新鮮な感じがした。姉に自慢したい（図 13）。

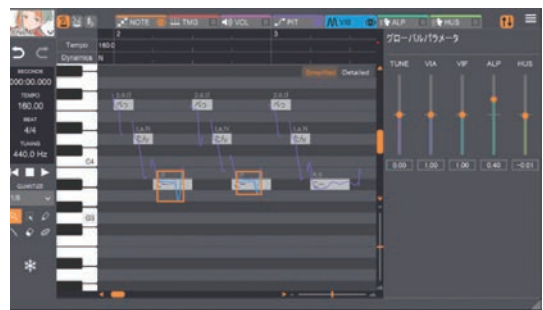


図 13 「VoiSona」(ボイソナ) (学生⑭)

4.4. 音楽合成「CoeFont STUDIO」

名称	URL
「CoeFont STUDIO」	https://coefont.cloud/

AIにおけるディープラーニング（深層学習）を活用することで、人間の声を再現して Web 上で音声合成が出来る。Web 上でアクセント調整やスピード調整などの機能を利用し、作成した音声はダウンロード可能である。キャラクターの2つの声を選択できる。「アリアル」は芯の通った声質、「ミリアル」は透き通った声質が特徴である。まず「無料で始める」を押す。メールアドレスを入力するとメールされて来る。メールに名前とパスワードを入力すると使えるようになる。「アリアル」か「ミリアル」を選び「今すぐ利用」を押す。「声を聞いてみる」を押して実際の声を聞く。右側の「CoeFont を使う」を押し、好みに合わせ下の「untitled」を押すと、セリフ入力欄が出る。セリフ入力欄にプロンプト（テキスト）を入力して「テキスト編集を終了」を押す。

学生の感想⑮

「様々なセリフを AI でアナウンサーや声優に読ませることが出来た。選んだのはアリアルで、芯の通った声の特徴的である。喜怒哀楽を変更できるので、様々な感情で喋らせることが出来る。アクセントも変更できるため自分に合った喋り方で設定可能であった。私は将来のエンタメを見据え AI についてもっと勉強したいと思った。」（図 14）



図 14 「CoeFont STUDIO」(学生⑮)

学生の感想⑯

「ティックトックやインスタのショートムービーに流れる声は本 AI が使われていると知れた。「アリアル」の声は女性だが男性的部分があったので、かわいらしい声の「ミリアル」を選択した。読み上げてもらった時にアクセントがおかし

い箇所は直すことが出来るので、聞き取り易くすることが出来た。動画配信する際に自分の声を出さず AI に話してもらえるため、喋りが苦手な人や活舌が悪い人には便利と思った。」（図 15）



図 15 「CoeFont STUDIO」(学生⑯)

学生の感想⑰

ミリアルの方が高い声で、優しいの会話が出来ると思いミリアルを選びました。「う」というところが「ー」になることにより砕けた会話になります。一度「う」に直して聞いてみたのですが、少し機械が話している感じが強調されてしまったので、普段の会話のように音声を作りたい時は「ー」にした方が良いと思いました。（図 16）



図 16 「CoeFont STUDIO」(学生⑰)

学生の感想⑱

「ミリアルに好きな Mrs.GREEN APPLE の新曲「Soranjī」の歌詞を歌わせました。優しさや神秘さのある曲の雰囲気合うのはミリアルであると感じミリアルを選びました。今後、AI による歌が身近になるので、どのように進歩するか楽しみです。」（図 17）



図 17 「CoeFont STUDIO」(学生⑬)

5. アンビート (Am Beat)

<https://tollite.yamaha.com/Ambeat/>

スマホで撮影した動画に付けるオリジナル音楽を作曲できる AI である。音楽や動画からどのような感想を抱くかを AI に学習させ、画像や動画、文章を入力すると、シーンに合う BGM を提案する。動画や写真のデータをアプリ上にアップロードするだけで、そのビジュアルコンテンツのイメージやシチュエーションにマッチした音楽を自動で作成できる。自動作成された BGM を、アプリ画面での簡単な操作で自由にアレンジすることも可能である。音楽の専門知識がなくても、テンポを変更したりコード進行や音楽の構成を選んだりして好みに合わせアレンジ出来る。BGM に歌を付加することも可能である。画面上で歌詞を入力すると、歌詞に合わせてボーカロイドの歌声による歌唱が生成される。毎月 10 本まで無料である。生成音楽は、商用または非商用を問わず BGM として利用できる。ユーザーは自身の動画に使うことが出来る。

6. 「Stable Audio」

<https://www.stableaudio.com/>

2023 年 9 月、「Stability AI」は、生成 AI によってテキストから音楽・サウンドを生成する「Stable Audio」を発表した。無料で 45 秒のトラックが生成できるサービスに加え、商用利用可能で 90 秒のトラックを生成する「Pro」サブスクリプションも用意された。プロンプトを入力する

と、希望するオーディオの長さに合わせて音楽が自動生成される。技術的には、画像生成と同様に「潜在拡散モデル (latent diffusion model)」を使用している (図 18)。45 秒なら無料であるため、学生にも実習してもらった。

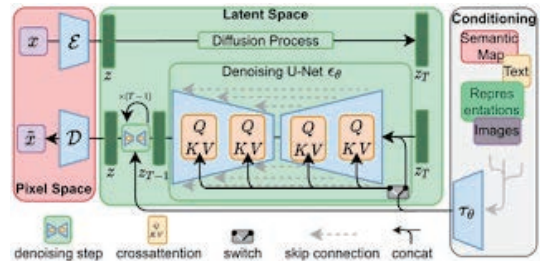


図 18 潜在拡散モデル (latent diffusion model)

出所: Louis Bouchard (2022) "How Stable Diffusion works? Latent Diffusion Models Explained A High-Resolution Image Synthesis Architecture: Latent Diffusion"

7. まとめ

画像や動画を生成する「生成 AI」が盛り上がる一方、音楽生成に関する研究は少し遅れている。忠実度の高い音楽生成を行うためには、様々なスケールで複雑な信号やパターンをモデリングする必要があることが理由である。音楽は、一組の音符から複数の楽器を含む大域的な音楽構造まで、局所的かつ長距離的なパターンで構成されるため、生成することが非常に難しい分野である。ゼミナールや演習実習では、生成が難しいとされる「音楽」を AI で作る実習を行った。AI の発達により、個人が様々なデジタルツールを使ってオリジナルを作って提供する時代に移ろうとしている。プロの仕事であった動画でも小説でも広告で



図 19 カンデオホテルズ六本木



図 20 カンデオホテルズ上野公園

も、もはや簡単に作ることが出来る。プロの作品を受動的に受け取るのではなく、自分でクリエイティブした方がライブ感を高く盛り上げられる。そして達成感が得られる。チームで意見が違うメンバーに付度する煩わしさからも解放される。

読者や視聴者など鑑賞側にとっては、人間が作ったかは関係ない。感動できれば、面白ければ、心地良ければ良い。石鍋 (2023) は「万葉集には『詠み人知らず』の歌も多い。近代以前の作者不明の工芸や美術にも人は感動する」と言う⁽⁸⁾。カンデオホテルズは館内のBGMをAI「アマデウスコード (Amadeus Code)」が生成したオリジナル曲を流すが、2023年7月26日、筆者 (植田) が六本木 (図 19) で8月1日に上野公園 (図 20) で朝食バイキングを取りながらアマデウスコード生成のBGMを聴いた。人間の作曲か、AI生成曲か⁽⁹⁾、気にも留めなかった⁽¹⁰⁾。アマデウスコードは学生に実習してもらった「サウンドロー」(3.1 項) の競合相手であるが、有料であるため学生には試してもらっていない。

名称	URL
「アマデウスコード」	https://amadeuscode.com/

名称	URL
「カンデオホテルズ」	https://www.candehotels.com/ja/

カンデオホテルズはビジネスホテルの「3つ星」、高級ホテルの「5つ星」の中間、「4つ星」を標榜する。「ここでしか出来ない体験」を目指すため、

自動作曲 AI「アマデウスコード」を使い、ホテルで使うBGM約100曲を作成した。「朝食の時は前向きで元気な雰囲気、深夜のチェックインは深く静かな感じで」というプロンプトで依頼した。季節やグレードに合わせ何曲でも作れる。他のどこにも流れないオリジナル曲にホテル側は「非常に満足している」。「アマデウスコード」の井上純 CEO は、グローバルチャート上位10位に食い込んだ過去60～70年間のヒット曲の他、定番のクラシックやジャズ1,300曲以上をAIに学習させた。

観賞側はアーティストの生活を考えてお金を払うのではなく、自分が満足したかどうかで「支払意思額」(WTP:「willingness to pay」)を決定する。そういう意味では、AI生成音楽にお金を払おうとは思わない。慶應義塾大学の今井むつみ教授 (2023) は「ホテルのロビーにある自動ピアノは心地良くてコーヒーを飲みながら聴くには良いが、お金を払って自動ピアノのコンサートを聴きに行く人はいない」と指摘する⁽¹¹⁾。「ホテルグランバッハ東京銀座」2階フロントロビーにあるヤマハ自動演奏ピアノは無人演奏する (図 21)。

ホテルやレストランにおいて無料で聴くBGMはAI生成音楽で構わないが、価値ある音楽に対しては高額を支払っても聴こうとする。テイラー・スウィフトのワールドツアーは各地でインフレ効果を巻き起こした。プロフェッショナルは多額のお金を払ってもらえる価値ある者とタダでしか聴かれずAIに代替される者に「二極化」される。「人間が作った音楽とAIが作った音楽、どちらが良いですか?」という問いは妥当ではない。タダ



図21 「ホテルグランパッハ東京銀座」自動演奏ピアノ

なら人間の方が良いが、大半の人間はタダで提供せずお金を請求するからである。何れにしても、音楽生成AIの成否は、実際に音楽を生成してみた人にしか分からないというのが結論である。私たちが可能性を持つ、新しい技術の誕生を目の当たりにしていることは間違いない。

音楽生成AIを用いた実習のポイントは「**創作の楽しさ**」が学生の学習に対するモチベーション社会にもたらす効果である。オリジナル音楽を作って発表すること自体がまず楽しい。そして「苦しみながら頑張った」よりも、楽しみつつ何かオリジナルの創作物ができる実感を持てることが大切である。生成AIをきっかけに多くの学生に創作の楽しみを知ってもらい、結果として「鑑賞」するだけに過ぎなかった受動的学びから、「製作（参加）」する能動的学びに変えることで、メディアやエンターテインメントを学習するモチベーションが高まれば良い。音楽はAIによって生まれるのではなく、好奇心や創造性のある人間によって作られる。最も重要な要素はAIそのものではなく人間である。生成AIの登場により大学も評論家的な受動教育から、自らが音楽をクリエイトする能動教育へと進化できる。

8. 今後の展望～AI芸術の先駆と拡張

今は欲しいモノが直ぐにネットで手に入る便利な社会である。工夫する必要がなく創造性が要らない社会でもある。学生のうちにもっと工夫する機会があって良い。生成AIはそのチャンスを与

えてくれる。人間が歌えないボカロ曲を聴いて育った「Ado」や「YOASOBI」は高音で高速、プレスなしという超人的な楽曲が生み出した。YOASOBIのボーカル「ikura」は、ボカロのラインを努力により習得して人間の声で歌うことを可能とした。AIでは出せない切なさをきちんと出すため、言語を超えて世界に伝わった。人間がピアノで弾くために人間が作る曲の難易度は、10本の指で演奏できる範囲、メロディーは人間が美しいと感じる範囲に留まる。AIや自動演奏ピアノであれば、演奏者としても人間の枠を超越できる。このように「上位層」のレベルは大幅に向上しており、「中間層」は相当頑張らないと追いつけない。カラオケでの歌唱や楽器演奏が容易になるべく「中間層」に受けるヒット曲は近年、ユニゾン（ハーモニーなし）、演歌も「こぶし」なし、ギターも3つのコードで演奏できるなど、低水準の楽曲が増えていた。音楽界の質を高めるためには、「中間層」の底上げを必要とする。

テレビや新聞で報道されるのはChatGPTが中心であるため、「生成AI＝テキスト」というイメージを持っている人が少なくない。生成AIによって誰もがクリエイターになれる時代になった。一部の職種ではハードルが高かったが、音楽アーティストや動画クリエイターになれる時代となった。AIによる音声合成の技術は既に人間のような自然な声を合成できる。そこに留まらず、アニメ声のような感情溢れる色々な音声も真似できる。合成音声のクリップの一部を変更して元の音声ストリームにもう一回挿入することも出来る。課題は、生成AIの能力と言うよりも、AIに指示を出す人間の想像力がAIの出す結果に追いつかない点である。どのような表現で見せたいか、同じテーマをどのような切り口で考えることを考える人間の想像力の方が全く追いついていない。AIが作詞作曲してくれるとしても、指示を出す人間の能力の限界が常に足枷となる。AIが人間の想像力を大きく増強することは確実である。

9. 今後の展望～音楽生成 AI の可能性

現在の「AI 芸術」はあくまでも人間が AI を道具として使って作ったもので、人間の美意識の範囲に留まる。現時点で AI が自ら創作する段階には到底至っていないが、将来もし自意識が生まれることがあれば、人間の美学を超えた美術が生まれる。AI が次に向かうステージは、「AI と AI がつながる世界」である。「AI 対 AI」「AI 同士のやり取り」が考えられる。今後の AI は人間の作った過去作を当てにせず、自立して創作する⁽¹²⁾。将棋 AI は、初期は過去の名人の棋譜を参考にしたが、途中から棋譜を学習せず AI 同士で 24 時間 365 日対局して強くなった。AI が人間に学ぶのではなく、藤井棋士が AI に新手を学ぶようになった。同じことが音楽や映像でも起こるため、現在、文句を言っている人たちが後年、AI に学ぶ時代になる。金融分野では、株の売買などアルゴリズムが支配的になっており、AI 同士のやり取りになっている。NTT の川添雄彦 CTO (2023) は「これまでの考え方が根底から覆る可能性があります。現在は、AI の相手が人間なので、人が分かり易い形にデータを変換しています。しかし、AI 同士がやりとりするようになると、人間には理解できないが、AI にとってはより最適な処理のやり方が出て来ると考えられます。その時、今までの技術や理論は全て崩壊します」と言う⁽¹³⁾。残念ながら、時代を変えるイノベーションを起こす学生に対して教員の態度は冷たい。自らの既得権益を守るため、自分の掌の中で動く学生を好む傾向にある。日本の大学が、個人の資質だけでなく「挑戦」や「異端」を受け入れて欲しいと願う。作曲家コンロン・ナンカロウ氏は、AI の誕生前から人間の限界を超えた演奏を可能とした自動演奏ピアノのための作曲を多数生み出した（図 22）。そのような作品を現代芸術家たちは「AI 芸術の先駆」と位置付ける。「機械の演奏を前提とするナンカロウ氏の曲は音楽の概念を拡張する。鑑賞する側も試される。この点は AI が作る音楽も同様である」と指摘される⁽¹⁴⁾。



図 22 コンロン・ナンカロウ

ここまで、音楽生成を技術や演出という観点で、専門ゼミナールや演習実習における現場からアイデアや事例を紹介した。生成 AI の優れた点は「AI の大衆化」を実現させた点である。音楽のように正解が曖昧で色々な解釈が出来る分野は、AI によって知性と感性を拡張できる。生成 AI の活用が進むには、若者に分かり易く伝えたり、実習を通してアドバイスしたりする作業が必要となる。実際に作ってみると想像を超えた生成物が手に入る事が分かる。普及のカギは「人材」である。先端科学に関わる人の裾野を広げることが、日本を元気づかせる。日本は人口が減り人材が先細ることが間違いないため、先端分野にしっかり取り組まなければいけない。AI 教育が行われる上で、これらの現場の実践的な知見がヒントになれば幸いである。

謝辞

本稿の作成に際し 2022 年度及び 2023 年度「卒業研究」「専門ゼミ」「演習 C」「実習 C」受講生の協力を得た。この場を借りてお礼申し上げたい。但し、本稿に関する誤りは筆者に帰属している。

参考文献

- [1] Louis Bouchard (2022) "How Stable Diffusion works? Latent Diffusion Models Explained A High-Resolution Image Synthesis Architecture: Latent Diffusion"
- [2] 人工知能学会誌『人工知能 38 巻 4 号 (2023 年 7 月)』
- [3] 木石岳 (2023) 『歌詞のサウンドテクスチャー』白水社

《注》

- (1) 山口一郎 (2023), 2023年4月16日付日経MJ14面
- (2) 発音の修正を担当したイ・ヒョン氏の所属事務所「HYBE」によると、AIを使ってイ・ヒョンさんの歌声から息づかいなどを残したまま発音データのみを抽出し、ネイティブの発音データに合わせて修正を加えた。発音の修正を担当した「HYBE」は「AIは音楽界にとって脅威ではありませんか」という問いに対して、「危険でもあり、脅威でもあります。ただ、AIを含む技術は道具だと思います。AIが歌手の代わりに歌うことが出発点ではなく、歌手が選択できる多様なオプションの一つだと考えています。新たな技術が登場した時、議論が盛り上がり、その過程で社会的合意が生まれることになると思っています。そして今後は「今回の曲を公開したことで、関心が高まっています。他のアーティストから『世界の多くの徴収に直接、この言葉で聴かせたい』というニーズがあれば、いつでも検討します」と言う。2023年7月20日付朝日新聞9面
- (3) マイケル・サンデル (2023), 2023年6月30日付日本経済新聞7面
- (4) 2020年、美空ひばりの過去の歌唱データを学習する「ボーカロイドAI」による歌声合成技術でヤマハが「AI ひばり」を開発したところ、山下達郎は自身のラジオ番組で「一言で申し上げると冒険です」と言い切った。一方で、坂本龍馬が司馬遼太郎の小説によって維新の立役者としてより広く知られるようになったのと同様に、後世の者が故人の資料データを基に人物像を造形することは繰り返し行われて来たという指摘もあった。2020年3月14日読売新聞31面
- (5) 柴那典 (2023), 2023年11月24日付日経MJ4面
- (6) 木石岳 (2023)『歌詞のサウンドテクスチャー』白水社
- (7) BGMとして考えられるのがゲームであるが、ゲームBGMには難しい側面がある。映像に合わせて盛り上がり下がりを演出することは難しい。モンスターを恰好よく倒す音楽を流しても、プレイヤーがその通りにモンスターを倒せるか分からない。映画BGMと異なる点である。また、音楽の切り替わりによってゲームへの没入感を失うリスクも存在する。
- (8) 石鍋仁美 (2023), 2023年4月2日付日本経済新聞12面
- (9) 「鑑賞者は心潤うメロディーに心を預ける」と河原啓子 (2023) は言う。2023年8月9日付朝日新聞11面
- (10) 全国に25施設あるカンデオホテルズのフロントでは、季節や時間帯に応じたオリジナルのBGMを流している。どの曲もAIが作り出した曲である。Amadeus Code（アマデウスコード）が2019年から提供を始め、制作したオリジナルBGMは100曲を超える。2分前後の曲なら、AIが10万〜20万の計算を行い4秒以下で生成する。井上純CEOは旧ジャニーズ事務所にプロデューサーとして所属して「KAT-TUN」「嵐」の2グループを担当した。過去のヒット曲を分析したところ、「売れる曲の黄金比が見えてきた」。曲を構成する要素はメロディー・ハーモニー・リズムの3要素があるが、アマデウスではメロディーに特化した生成AIを開発した。楽譜をアルファベットと数字の羅列に置き換えたテキストデータを用いている。データは、1〜4音で構成されたメロディーの断片「Lick（リック）」として蓄積される。アマデウスコードが音楽生成の分野で世界に先行するのは、音楽は日本語の文の構造と近いからである。英語や中国語の場合、主語、動詞と来て、その後に目的語が続く。日本語は主語の後に複数の目的語が続き、最後に動詞がある。両者では、文章を生成していくアルゴリズムは大きく異なる。音楽は、出だしの音から最終的にケーデンス（終止）に向かってメロディーを連ねて行く。主語から始まって動詞で終わる日本語と似ている。「誰でもイメージさえあれば作曲できるという世界が我々のゴールだ」と井上氏は言う。AIと人間が力を合わせて、新しいものを生み出して行く狙いである。佐藤嘉彦 (2023)「日経ビジネス」オンライン (2023.5.1) <https://business.nikkei.com/atcl/gen/19/00537/042500019/>
- (11) 今井むつみ (2023), 日経クロステック『ChatGPT 産業革命』129p.
- (12) AI同士が自律的にやり取りする「オートGPT」や「AI エージェント」などが登場している。人間とAIがやり取りする場合、自然言語というインターフェースが必要であったが、AI同士のやり取りには必要なくなる。
- (13) 川添雄彦 (2023), 2023年7月19日付日経産業新聞16面
- (14) 佐堀万梨映 (2023), 2023年10月6日付日本経済新聞35面