

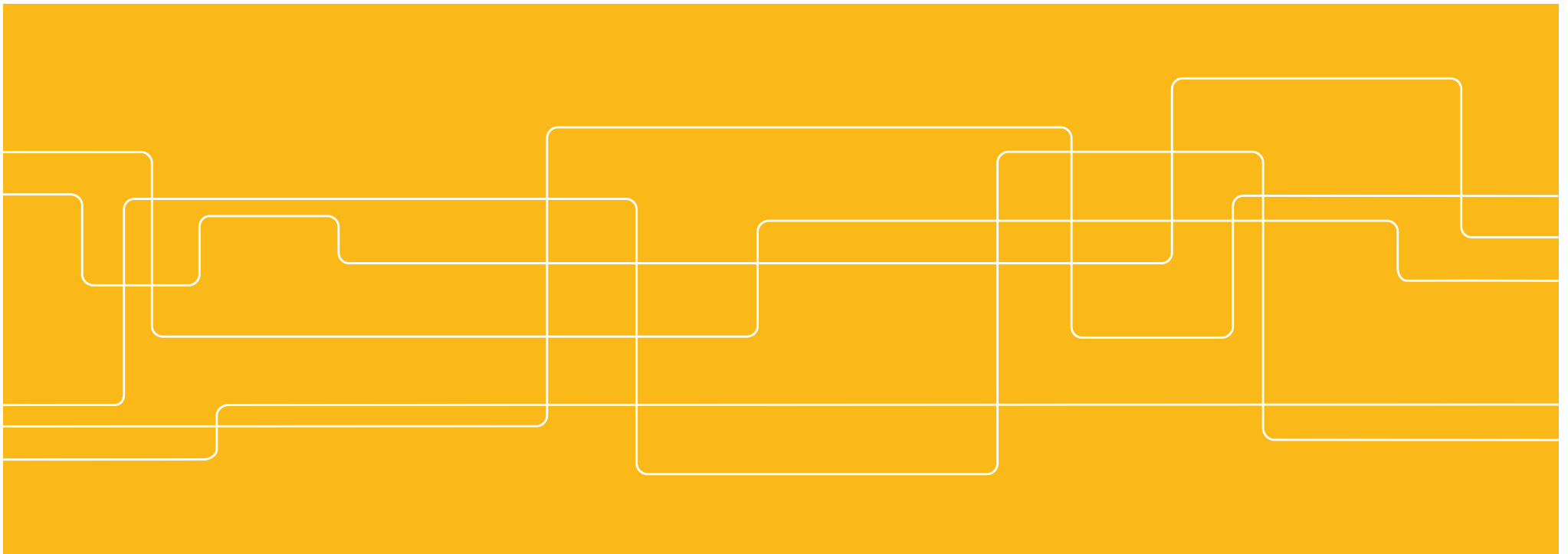


# Tunes from the Ai Frontiers

Bob L. T. Sturm

*Speech, Music and Hearing Division*

School of Electronic Engineering and Computer Science





## Machine learning is:

- Methods for making machines learn from data.
- One of the hottest areas of research in Ai today!
- Hugely successful due to a confluence of data, computational resources, advancements in efficient and effective algorithms, and massive amounts of capital.



## Folk music is:

- An historically motivated categorization of musical practice among “common” people sharing a regional identity
- Often accompanying other practices, such as dancing or story telling
- Passed aurally with provenance



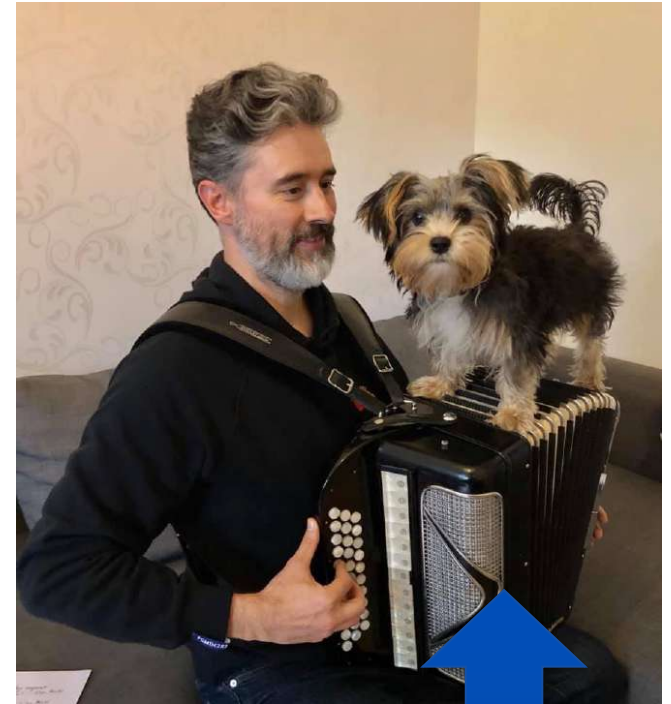
Johnny O'Leary



# Ai Folk?

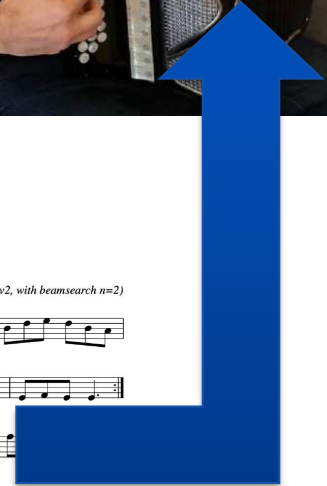
## Faux Folk?

### trAditional Music?



*folk-rnn (v2, with beamsearch n=2)*

A





# Machine Folk

**THE CONVERSATION**  
Academic rigour, journalistic flair

**'Machine folk' music composed by AI shows technology's creative side**

March 31, 2017 10.40am BST



**Bob Sturm**

Lecturer in Digital Media, Queen Mary University of London



**Oded Ben-Tal**

Senior Lecturer in Music Technology, Kingston University

<https://theconversation.com/machine-folk-music-composed-by-ai-shows-technologys-creative-side-74708>



## An Example: “The Boys of Ballinaburre”

<s> M:6/8 K:Edor | : B E E B E E | F E  
D D E F | G E E B E E | B d e d B A |  
B E E B E E | F E D D E F | G A B A F  
D | E F E E 3 : | | : B e e e f e | d B  
c d B A | B e e e d e | f d f d 2 e |  
B e e e f e | d B c d B A | G A B d B  
A | F E D E 3 : | </s>



# An Example: “The Boys of Ballinaburre”

*folk-rnn (v2, with beamsearch n=2)*



# Some background

*How did I get here?*







# The Unreasonable Effectiveness of Recurrent Neural Networks

May 21, 2015

There's something magical about Recurrent Neural Networks (RNNs). I still remember when I trained my first recurrent network for [Image Captioning](#). Within a few dozen minutes of training my first baby model (with rather arbitrarily-chosen hyperparameters) started to generate very nice looking descriptions of images that were on the edge of making sense. Sometimes the ratio of how simple your model is to the quality of the results you get out of it blows past your expectations, and this was one of those times. What made this result so shocking at the time was that the common wisdom was that RNNs were supposed to be difficult to train (with more experience I've in fact reached the opposite conclusion). Fast forward about a year: I'm training RNNs all the time and I've witnessed their power and robustness many times, and yet their magical outputs still find ways of amusing me. This post is about sharing some of that magic with you.

*We'll train RNNs to generate text character by character and ponder the question "how is that even possible?"*

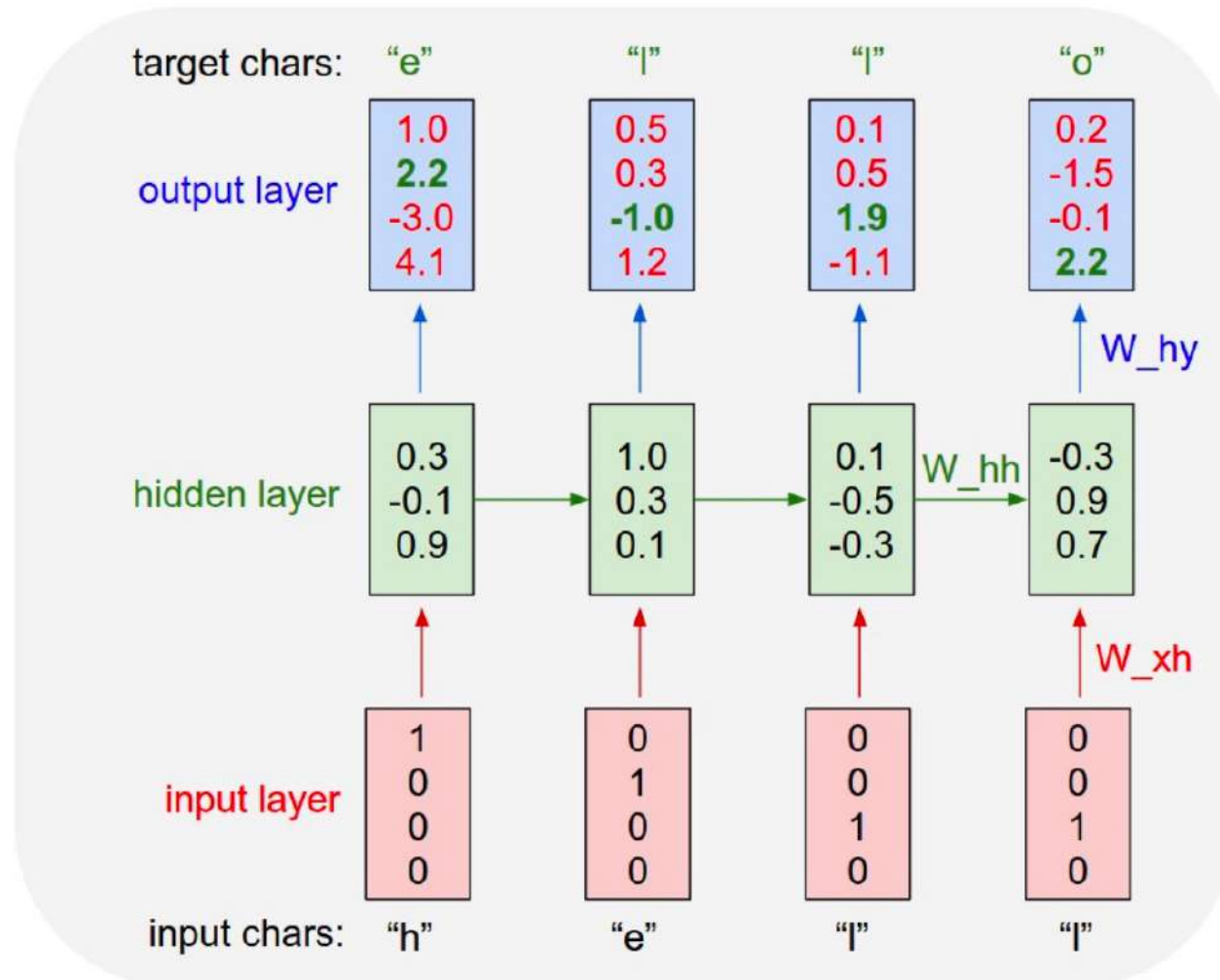
By the way, together with this post I am also releasing [code on Github](#) that allows you to train character-level language models based on multi-layer LSTMs. You give it a large chunk of text and it will learn to generate text like it one character at a time. You can also use it to reproduce my experiments below. But we're getting ahead of ourselves; What are RNNs anyway?

## Recurrent Neural Networks

Sequences. Depending on your background you might be wondering: *What makes Recurrent Networks so special? A glaring limitation of Vanilla Neural Networks (and also Convolutional Networks) is that their API is too*



# A basic recurrent network





PANDARUS:

Alas, I think he shall be come approached and the day  
When little srain would be attain'd into being never fed,  
And who is but a chain and subjects of his death,  
I should not sleep.

Second Senator: ... produced upon my soul,

```
<page>
  <title>Antichrist</title>
  <id>865</id>
  <revision>
    <id>15900676</id>
    <timestamp>2002-08-03T18:14:12Z</timestamp>
    <contributor>
      <username>Paris</username>
      <id>23</id>
    </contributor>
    <minor />
    <comment>Automated conversion</comment>
    <text xml:space="preserve">#REDIRECT [[Christianity]]</text>
  </revision>
</page>
```

was grounded  
], associated  
he Portugal  
Cantonese  
ng. The  
Kingdom  
], known  
flict.  
ce that  
nal



# Source Code

karpathy / char-rnn

Watch 569 Star 10k Fork 2.4k

Code Issues 87 Pull requests 22 Actions Projects 0 Wiki Security Insights

Multi-layer Recurrent Neural Networks (LSTM, GRU, RNN) for character-level language models in Torch

88 commits 1 branch 0 packages 0 releases 18 contributors

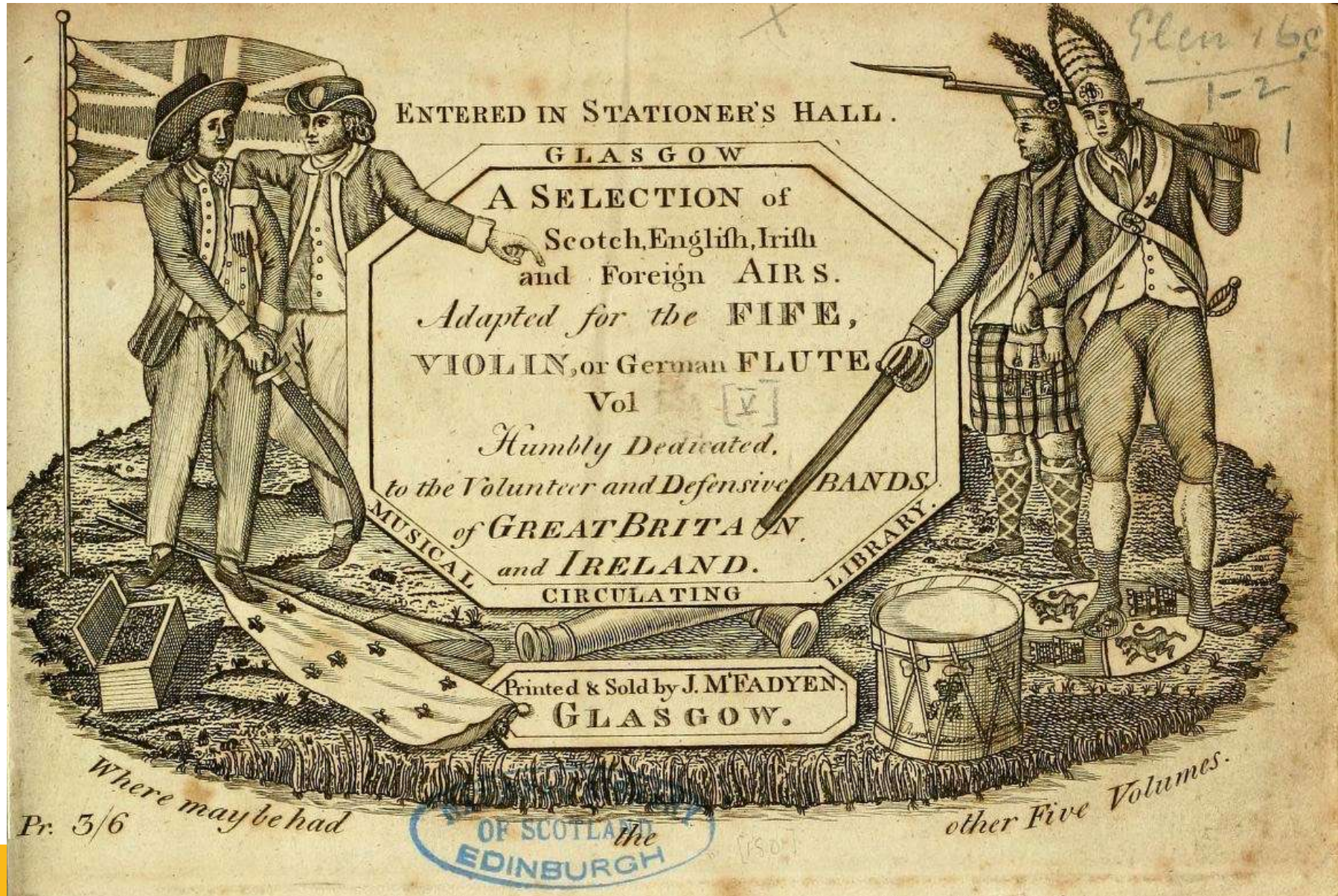
Branch: master New pull request Create new file Upload files Find file Clone or download

karpathy Merge pull request #164 from gdb/master Latest commit 6f9487a on Apr 30, 2016

data/tinyshakespeare	first commit	5 years ago
model	changing default LSTM initialization to use biases of 1.0 for the for...	5 years ago
util	Fix unclear errors	4 years ago
.gitignore	Add t7 files to .gitignore	5 years ago
Readme.md	Update Readme.md	4 years ago
convert_gpu_cpu_checkpoint.lua	fixing a bug introduced in previous commit. We have to use doubles no...	5 years ago
inspect_checkpoint.lua	add openc1 to sample.lua and inspect_checkpoint.lua, add link to clto...	5 years ago

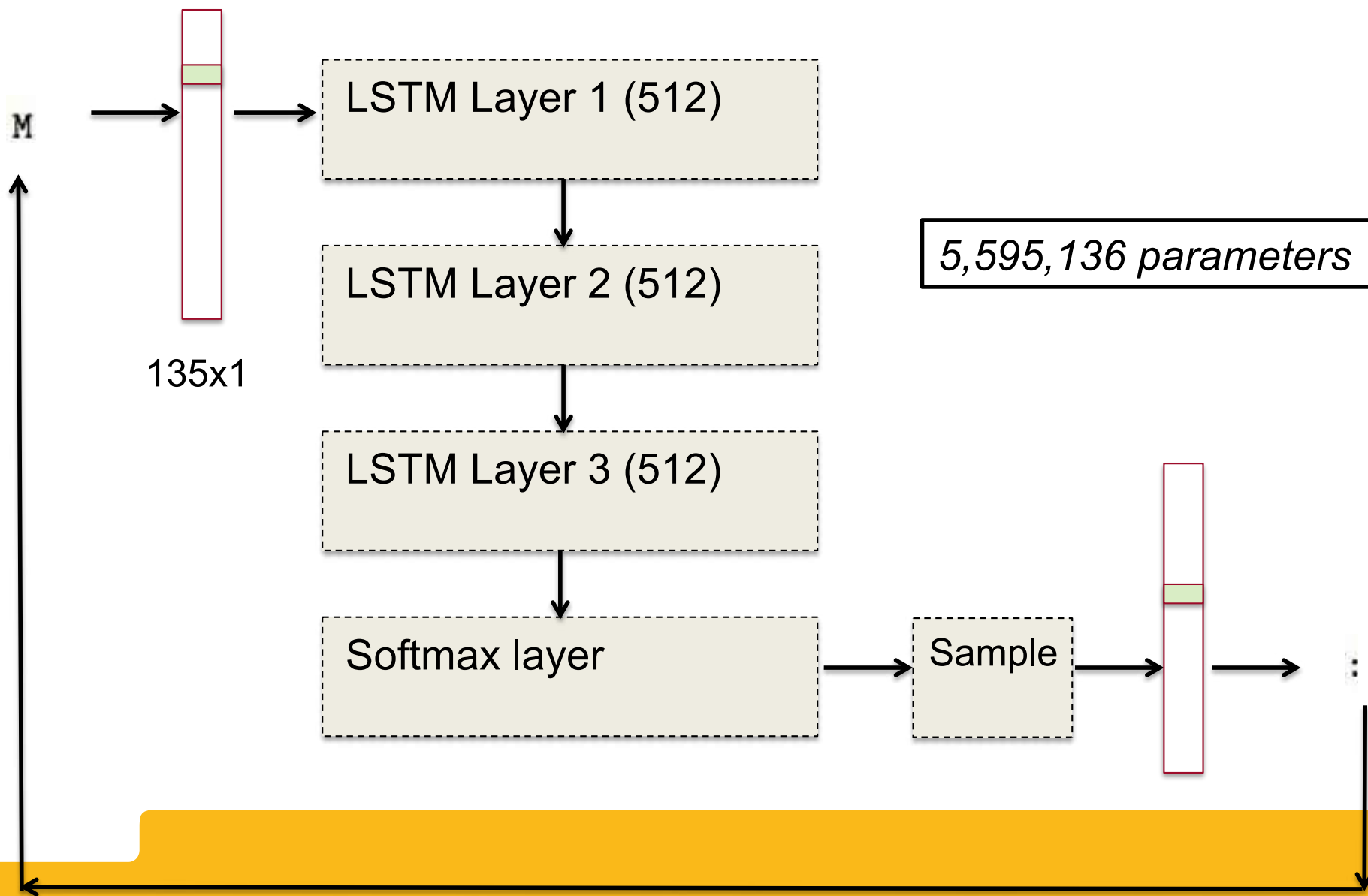
<https://github.com/karpathy/char-rnn>

# First experiments (1180 tunes)





# folk-rnn (v1): architecture





# Meh

## Lisl's Stis.

♩ = 120



## Quirch cathp'3b The Nille L' theys Lags Bollue's

♩ = 120

Let's use more data!



# <http://thesession.org>

The screenshot shows a web browser window with the URL <https://thesession.org>. The browser's address bar and tabs are visible at the top. Below the browser window is a dark red navigation bar with the following links: [Log in or Sign up](#), [TUNES](#), [RECORDINGS](#), [SESSIONS](#), [EVENTS](#), and [DISCUSSIONS](#). The main content area has a yellow background. On the left, there is a logo for 'THE SESSION' with a large 'S'. To the right, the word 'Fáilte' (Irish for 'Welcome') is displayed. Below this is a search bar with a dropdown menu set to 'The Session' and a 'SEARCH' button. Underneath the search bar is a section titled 'Recent activity' which lists two recent events: 'Cami Francesa Carre added [Lost In The Loop](#) to their tunebook.' (2 minutes ago) and 'emmdee left a comment on the discussion [FS: 2004 Eamonn Cotter 4 keyed flute.](#)' (10 minutes ago). A solid yellow bar is at the bottom of the page.





## Example transcription from *thesession.org*

27,27,"Drowsy Maggie","reel","4/4","Edorian",":|E2BE dEBE|E2BE AFDF|E2BE dEBE|BABc dAFD:| d2fd c2ec|defg afge|d2fd c2ec|BABc dAFA| d2fd c2ec| defg afge|afge fdec|BABc dAFD|", "2001-05-21 03:47:39", "Jeremy"



# Example transcription from *thesescience*

data — vi sessions\_data\_clean\_v1 — 91x25

T: Drowsy Maggie  
M: 4/4  
L: 1/8  
K: Edor  
~E2 BE dEBE | ~E2 BE AFDF | (3EGE BE dEBE | dDdB AFDF |  
~E2 BE dEBE | ~E2 BE AFDF | GFEF GABc | dcdB A2 FA |  
Be (3eee e2 eg | fd (3ddd d2 ef | g3 g g3 g | fa (3aaa afed |  
d2 (3bbb bage | fa (3aaa afed | ~B3 B BAFA | ~d3 B AFDF ||

T: Drowsy Maggie  
M: 4/4  
L: 1/8  
K: Edor  
"Em"E2 GE BE GE | "Em" E2 GB "D"AD FD | "Em" E2 GE BE GE | "G" BA Bc "D"dA FD |  
"Em"E2 GE BE GE | "Em" E2 GB "D"AD FD | "Em" E2 GE BE GE | "G" BA Bc "D"dA FA: |  
|:"D"d2 fd "A"c2 ec | "D" de fg "D"af "Em"ge | "D" d2 fd "A"c2 ec | "G" BA Bc "D"dA FA |  
"D"d2 fd "A"c2 ec | "D" de fg "D"af "Em"ge | "D" af"Em"ge "D"fd"Em"ec | "G" BA Bc "D"dA FD: | ]

T: Drowsy Maggie  
M: 4/4  
L: 1/8  
K: Bmin  
||B2 FB aBfB|B2 FB ecAc|B2 FB aBfB|fefg aecA|



201712 — vi ~/research/2017/201712/output\_folkrrnnv1.txt

T: Soup At The Compaire, The  
 M: 4/4  
 L: 1/8  
 K: Bmin  
 |:d2BF d2cd|e2ge cA(3cBA|d2df d2cd|e2fg aecA|  
 d2fd (3BcB dB|A2fd cAFA|G2Bd a2af|gece d4|  
 dcd fdc|(3Bcd cA G2ED|FABc dAFA|G/A/B AG FDDA|  
 d2fd g2fg|A2Bc egfe|ABcd eAce|Bd (3cde ffd:|

T: Drowsy Magr  
 M: 4/4  
 L: 1/8  
 K: Edor  
 ~E2 BE dEBE e2~e2  
 ~E2 BE dEB'|(3Bcd ec d2Af|(3aaa ag fada|a2da edfa|baba geeg:|  
 Be (3eee e fdc d2ef|~g3e ~f3d|edcA ABcd|egfe d3e|  
 d2 (3bbb fdc defg|e2Ac defg|afbg aedf|edcB AGFE:|

T: Drowsy T: Down Grannson Itell  
 M: 4/4 M: 4/4  
 L: 1/8 L: 1/8  
 K: Edo Gb||  
 "Em"E'  
 "Em"F B3 f 6|-  
 |:"D |g| 3t'dena, 1|| T. |- - e TTg| ^Aafm -e| F5mf 2  
 "D", Becdgc B3Bcd^^c|f2A e ss4  
 T: d2e>mdd/gge|>gdA Cleeeeddcd3| d2e2B2' -2||1:|:  
 M: |AB8 EG g3^mee24d de e |( ndGggg f2 -2/2| WAE FGAd3/ d3  
 L <d2||| 2G "AG fBBBB|/d3fd2afg fe|g{r233DAr3A3 c=dcc|B2 c a3 ec

A FD:|]



# Sorpike's Cat

*folk-rnn (v1) + Sturm*

The musical score is written in treble clef, 4/4 time, with a key signature of one sharp (F#). It consists of six staves of music. The first staff is a single line of music. The second staff contains a triplet of eighth notes and a first/second ending bracket. The third staff features a dotted quarter note with a fermata and a first/second ending bracket. The fourth staff continues with a first/second ending bracket and a triplet of eighth notes. The fifth staff has a triplet of eighth notes and a first/second ending bracket. The sixth staff contains several groups of eighth notes highlighted with pink shading and a first/second ending bracket.



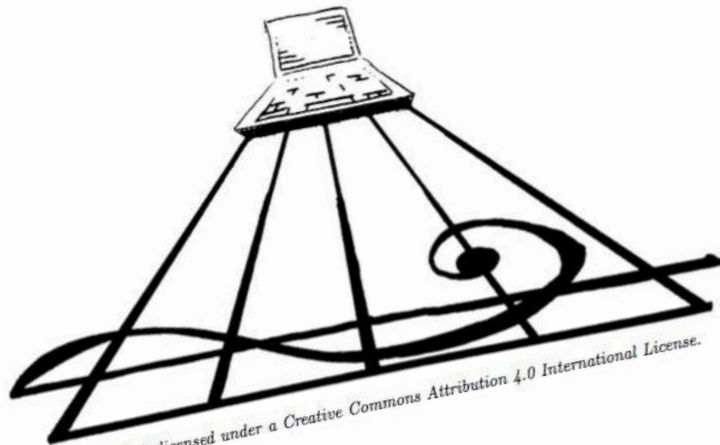
# 100,000 tunes in 34 volumes

The folk-rnn (v1) Session Book  
Volume 1 of 20\*



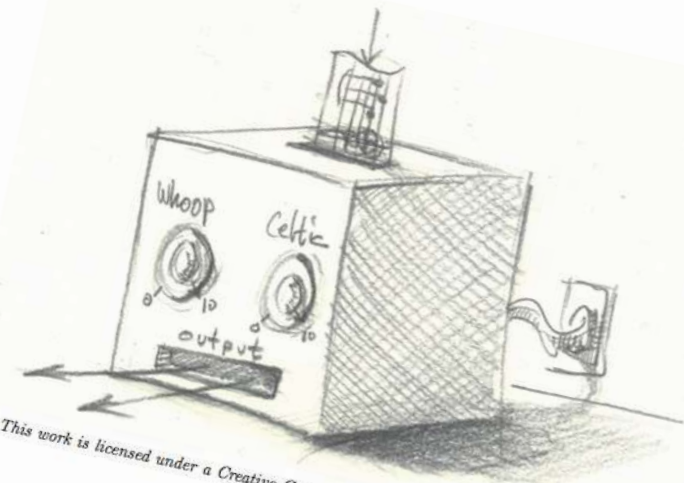
nder a Creative Commons Attribut

The folk-rnn (v3) Session Book  
Volume 1 of 4\*



This work is licensed under a Creative Commons Attribution 4.0 International License.

The folk-rnn (v2) Session Book  
Volume 1 of 10\*



This work is licensed under a Creative Commons Attribution 4.0 International License.



# The Endless Traditional Music Session (2015)

Music generated and titled by a recurrent neural network,  
trained on over 23,000 tunes from ABC code posted on The Session

This set is performed by *The First Me Schoast* (playlist updated every 5 minutes).



1. [Eld In The Weclanges](#)
2. [Pull Down The Battle](#)
3. [The Boos Of The Leaden](#)
4. [The Swan](#)
5. [The Rupnan Isopda](#)
6. [Ar Cashec A' Chan](#)
7. [The Clibbon's One](#)

Total number of tunes generated now: 35809

Created by: [Bob L. Sturm](#) and [João Felipe Santos](#). ([More info.](#))



# loquantur *RHYTHM*

presents



## *Folk-RNN*



Our artist this month is "**Folk-RNN**" - a computerized/ machine learning system "trained" on folk music that composes ORIGINAL music!

**Folk-RNN** was developed in London, England at Queen Mary and Kingston Universities and has composed over 35,000 original tunes – all based on

<https://soundcloud.com/sturmen-1/on-hold-millennial-whoop-reel>



# <http://thesession.org>

The screenshot shows a web browser window with the URL <https://thesession.org>. The browser's address bar and tabs are visible at the top. The website's navigation menu is dark red with white text, including "Log in or Sign up", "TUNES", "RECORDINGS", "SESSIONS", "EVENTS", and "DISCUSSIONS". The "DISCUSSIONS" link is highlighted with a yellow circle. Below the navigation bar is a yellow banner with the text "THE SESSION". The main content area has a light yellow background and features a "Fáilte" (Welcome) heading. Below this is a search bar with a dropdown menu set to "The Session" and a "SEARCH" button. The "Recent activity" section lists two items: "Cami Francesa Carre added Lost In The Loop to their tunebook." (2 minutes ago) and "emmdee left a comment on the discussion FS: 2004 Eamonn Cotter 4 keyed flute." (10 minutes ago). A solid yellow bar is at the bottom of the page.





# Comments @ thesession.org

## Re: On computer generated tunes

Jeez. A computer that noodles. That's all we need.

# Posted by *Mark M* 10 months ago.

## Re: On computer generated tunes

This has to be a windup, they're terrible!!

hahaha

# Posted by *irishfiddleCT* 10 months ago.

## Re: On computer generated \*sample\* tunes

Frankly when I listened to the samples the percussion was the only bit which may have redeeming value. The tunes, as they were, weren't. Tunes that is.

Basically it's crude turntabling without the sense of a musician familiar with the significance of various motifs & phrases.

# Posted by *AB* 10 months ago.

## Re: On computer generated tunes

Teach it to dance first?! 🤨

# Posted by *ceolachan* 10 months ago.

“I think it's reckless to send 3,000 machine-created [tunes] into the world.” -- *thesession.org* user Ergo

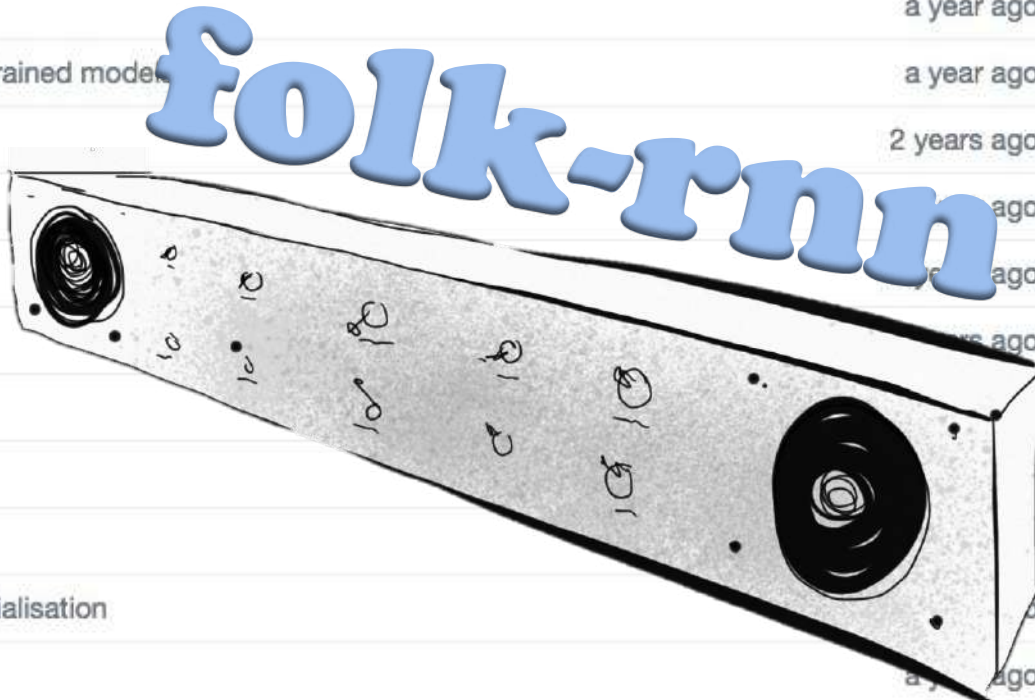
<https://thesession.org/discussions/37800>



# A vanilla LSTM with a special vocabulary

<https://github.com/IraKorshunova/folk-rnn>

boblsturm committed on GitHub Update README.md		Latest commit 52a7d37 22 hours ago
configurations	bugfix	a year ago
data	change to readme	a year ago
metadata	Updated metadata with newer trained model	a year ago
samples	add metadata and samples	2 years ago
soundexamples	Update README.md	ago
.gitignore	clean before checkout	ago
LICENSE	license	ago
README.md	Update README.md	ago
data_iter.py	added 1hot option	ago
logger.py	bugfix	ago
sample_rnn.py	added terminal output, fixed initialisation	ago
train_rnn.py	bugfix	ago




Sturm, et al., “Music transcription modelling and composition using deep learning,”  
in *Proc. Conf. Computer Simulation of Musical Creativity*, 2016.




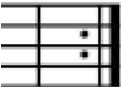
# Token types in the *folk-rnn* (v2) vocabulary


*Aim: design a vocabulary that has little ambiguity, and efficiently represents transcriptions.*


1. Transcription (2) <s>, </s>


2. Meter (7) M:4/4 

3. Mode (4) K:Cdor 

4. Measure (5) |:| 

5. Pitch (85) A c ^f 

6. Duration (25) 2 

7. Grouping (9) (3 [ ] 

- All titles, ornamentations, etc. removed
- All tunes transposed to root C
- 137 tokens in total



# Example transcription from *thesession.org*

27,27,"Drowsy Maggie","reel","4/4","Edorian",":|:E2BE dEBE|E2BE AFDF|E2BE dEBE|BABc dAFD:| d2fd c2ec|defg afge|d2fd c2ec|BABc dAFA| d2fd c2ec| defg afge|afge fdec|BABc dAFD|", "2001-05-21 03:47:39", "Jeremy"

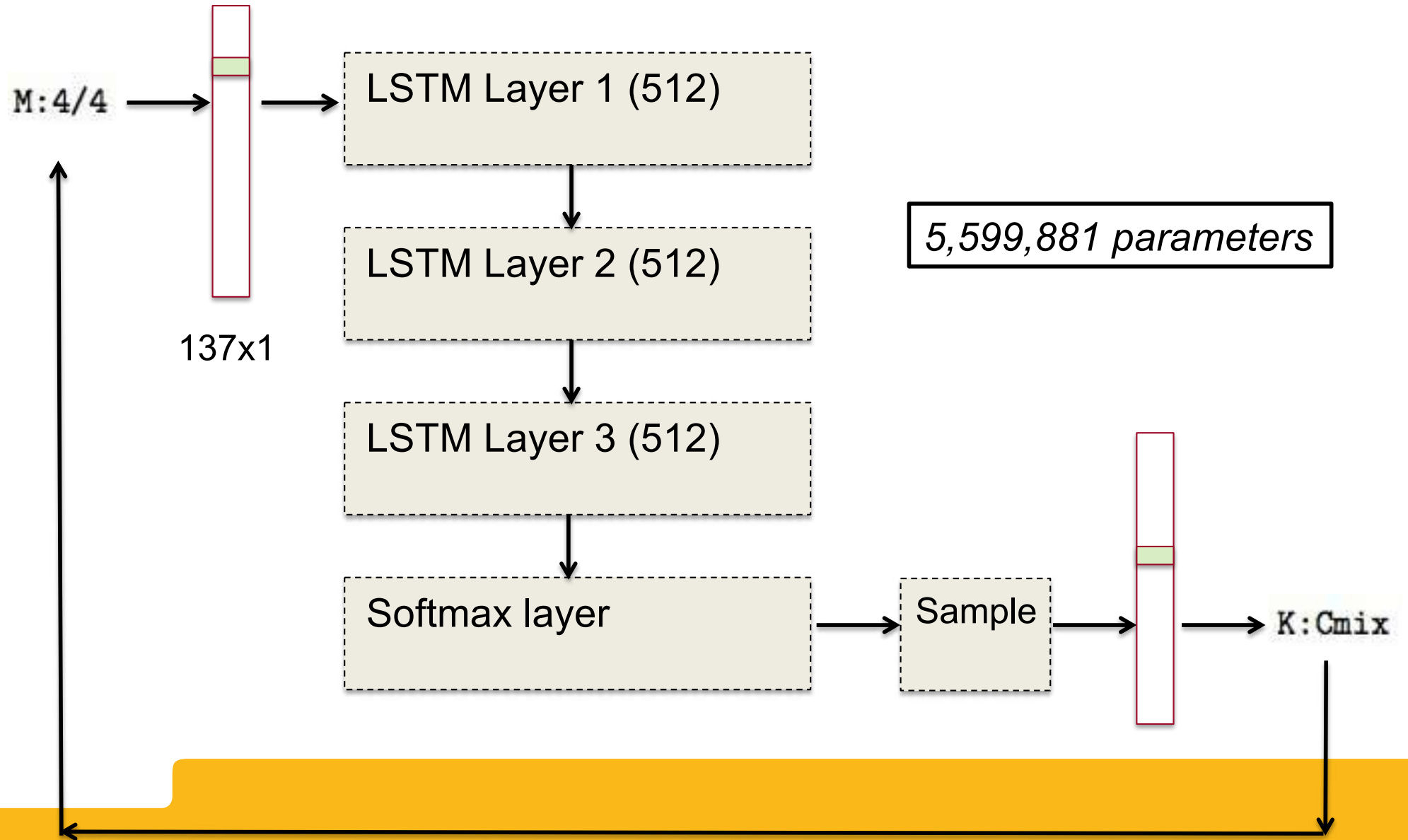


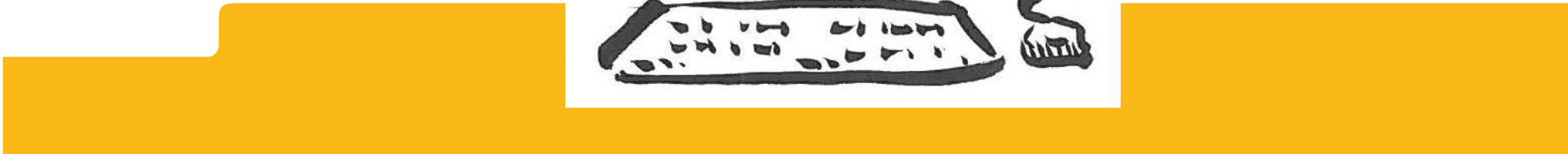
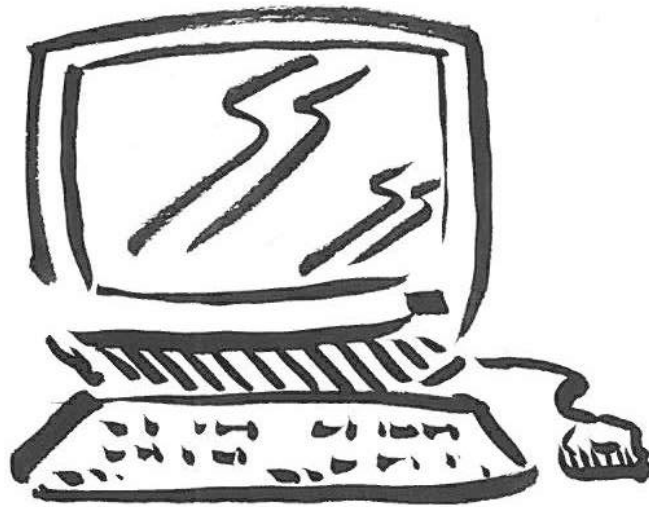
*extract, transpose, tokenise*

M:4/4 K:Cdor |: C2 G C B C G C | C2 G C F D B, D | C2 G C B C G C | G F G A  
B F D B, :| B2 d B A2 c A | B c d e f d e c | B2 d B A2 c A | G F G A B F  
D F | B2 d B A2 c A | B c d e f d e c | f d e c d B c A | G F G A B F D B, |



# folk-rnn (v2): architecture







# Human-in-the-loop composition

## Deep learning for assisting the process of music composition (part 1)

Edit

Posted on August 11, 2015 by Bob L. Sturm

*This is part 1 of my explorations of using deep learning for assisting the process of music composition. In this part, I look at some almost-winning output of a model trained by deep learning methods on over 23,000 folk tunes, and make improvements to produce a session-ready piece.*

## Deep learning for assisting the process of music composition (part 4)

Edit

Posted on August 15, 2015 by Bob L. Sturm

*This is part 4 of my explorations of using deep learning for assisting the process of music composition. In this part, I look at the process of composition using a model built using deep learning methods on over 23,000 folk tunes. Part 1 is [here](#). Part 2 is [here](#). Part 3 is [here](#).*

<https://folkrrn.org/>

**folkRNN**  
generate a folk tune with a  
recurrent neural network

PRESS TO GENERATE TUNE

Compose

MODEL

thesession.org (w/ :l

TEMPERATURE

1

SEED

359615

METER

4/4

MODE

C Major

INITIAL ABC

Enter start of tune in  
ABC notation

## FOLK RNN TUNE №1580

X:1580

M:4/4

K:Cmaj

```
cGGE GFEG|(3CCCD EGEG|cAdc ecdB|ACDE G3A|
(3cccCG EGGC|EDCD (3EEEG2|cGGG e2dc|(3ABcdB c3d:|
|:e3d cege|f3d Bdga|gfeg ecde|fegc A2GF|
EGG2 cdeg|(3fffaf dafd|e2eg fdBc|dedB c3d:|
```

The RNN properties were *thesession\_with\_repeats* with seed **441885** and temperature **1**.

The prime tokens were *M:4/4 K:Cmaj*.

Generated on **14/06/2018, 14:46:35**.

HEAR IT

Audio player interface showing a play button, a progress bar at 0:00, and a "Download MIDI" button.

SEE IT

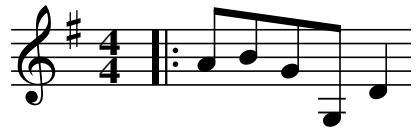
Musical score for the generated folk tune, displayed in four staves of treble clef notation. The score includes triplets and repeat signs.







# Theme from “Close Encounters of the 3<sup>rd</sup> Kind”



<s> M:4/4 K:Cmaj d e c C G 2



# Theme from “Close Encounters of the 3<sup>rd</sup> Kind”

*folk-rnn (v2) + Sturm*

The image displays a musical score for the theme from "Close Encounters of the 3rd Kind". The score is written in G major (one sharp) and 4/4 time. It consists of six staves of music. The first staff begins with a double bar line and a repeat sign, followed by a green highlight over the first four measures. The second and fourth staves include first and second endings, indicated by brackets and the numbers '1' and '2'. The fifth and sixth staves feature pink highlights under various measures. The music is primarily composed of eighth and sixteenth notes, with some rests and a key signature change to F major in the fifth staff.

## Two Short Pieces with an Interlude

1. March to the Mainframe
2. Interlude
3. The Humours of Time Pigeon



## Chicken Bits and Bits and Bobs



f1  
Bl  
P  
v  
c  
t



Sturmen

We three layers o' hidd'n units are



Sturmen

It Came Out From a Pretrained Net



3 years ago

#neural network

2 years ago



# “Bastard Tunes” by Oded Ben-Tal (2017)



Bastard Tunes - 1st movement

1,730 views · May 25, 2017

👍 14    🗨️ 0    ➦ SHARE    ≡+ SAVE    ...



**The Bottomless Tune Box**  
96 subscribers

SUBSCRIBED



<https://youtu.be/YZ2jb0ksOm4>



# Failure can be very interesting

Eight short outputs generated by a long short-term memory network with three fully connected hidden layers of 512 units each trained on over 23,000 ABC transcriptions of session music (Irish, English, etc.), and arranged by my own "personal" neural network trained on who knows what for who knows how long (I can't remember any of the settings)

0:00 / 9:46

Analytics Video Manager

Eight Short Outputs ... <https://youtu.be/RaO4HpM07hE>



“The Humours of Time Pigeon”

Arranged for ensemble:

<https://bit.ly/386wrnG>

Daren Banarsë



Jennifer Walshe



John Hughes



Laura Agnusdei



Paudie O'Connor  
Aoife Ní Chaoimh



Úna Monaghan

Torbjörn Hultmark



Luca Turchet



Zoë Gorman



Jennikel Andersson



# Music in the age of Artificial Creation

November 20 2017 7 PM  
St Dunstan and All Saints, Stepney High  
Street E1 0NR



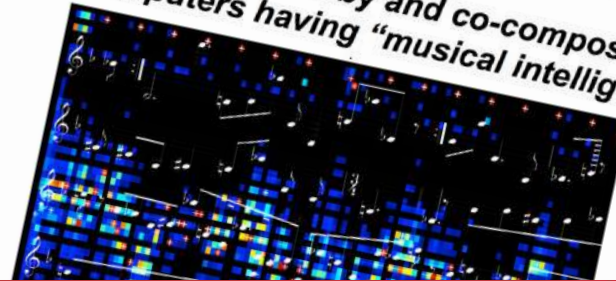
PRESENTED BY O'REILLY AND INTEL AI

## Artificial Intelligence CONFERENCE

BEIJING • NEW YORK • SAN FRANCISCO • LONDON

# PARTNERSHIPS

Music composed by and co-composed  
with computers having "musical intelligence"



### PUT AI TO WORK

8-9 OCT 2018: TRAINING

9-11 OCT 2018: TUTORIALS & CONFERENCE

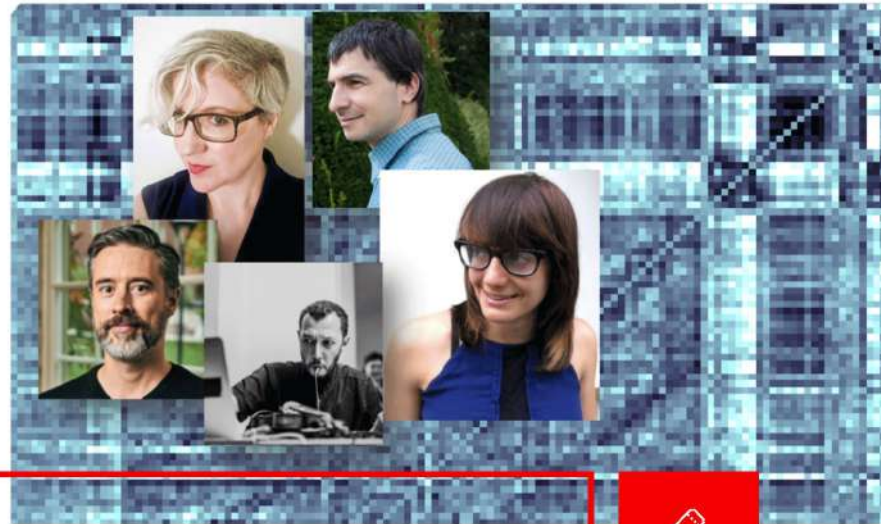
LONDON, UK

## MUME 2017 CONCERT

The Musical Metacreation Concert at the  
Eighth International Conference on Computational Creativity,  
Atlanta, Georgia

Monday June 19th. Doors 7pm. Music 8pm.  
Mammal Gallery, Atlanta.

Tickets \$20.



### Musik/Data – Data/Musik

🎵 Klassiskt, Elektroakustisk musik

🕒 tisdag 15 oktober 20.00 - 21.30



BOKA BILJETT



<p><b><i>Pieces for organ: The Glas Herry Comment &amp; X:7153</i></b> by folk-rnn + Deep Bach (2017)</p>	<p><b>Richard Salmon</b> <i>organ</i></p>
<p><b>Traditional Irish Sets</b> (with <i>folk-rnn</i> tunes in <i>italics</i>)</p> <ul style="list-style-type: none"><li>• <b>Jigs</b> (The Cuil Aodha, The Dusty Windowsill, <i>The Glas Herry Comment</i>)</li><li>• <b>Slow Reels</b> (Maghera Mountain, X:2897)</li><li>• <b>Fast Reels</b> (The Rookery, X:1068, Toss The Feathers II)</li></ul>	<p><b>Daren Banarsë and Musicians</b></p>
<p><b><i>March to the Mainframe, Interlude, The Humours of Time Pigeon</i></b> by Bob L. Sturm + <i>folk-rnn</i> (2017)</p>	<p><b>Ensemble x.y</b></p>
<p><b><i>Ed SheerAI vs XenAkls vs Aldele</i></b> by Nick Collins (2017)</p>	<p><b>Ensemble x.y</b></p>
<p><b>3 morphed pieces from “A Little Notebook for Anna Magdalena”</b> by J. S. Bach (1722) + MorpheuS (2017) <b>3 morphed pieces from “30 and 24 Pieces for Children”</b> by Kabalevsky (1937) + MorpheuS (2017)</p>	<p><b>Elaine Chew</b> <i>piano</i></p>
<p><b><i>Safe Houses</i></b> by Úna Monaghan + <i>folk-rnn</i> (2017)</p>	<p><b>Úna Monaghan</b></p>
<p><b><i>The Choice</i></b> by Úna Monaghan (2015) <b><i>The Chinwag</i></b> by Úna Monaghan (2015)</p>	<p><i>Irish harp, concertina, electronics</i></p>
<p><b><i>Pieces for organ: X:633 &amp; The Drunken Pint</i></b> by folk-rnn + Deep Bach (2017)</p>	<p><b>Richard Salmon</b> <i>organ</i></p>
<p><b><i>Chicken Bits and Bits and Bobs</i></b> by Bob L. Sturm + <i>folk-rnn</i> (2017)</p>	<p><b>Ensemble x.y</b></p>
<p><b><i>Bastard Tunes</i></b> by Oded Ben-Tal + <i>folk-rnn</i> (2017)</p>	<p><b>Ensemble x.y</b></p>

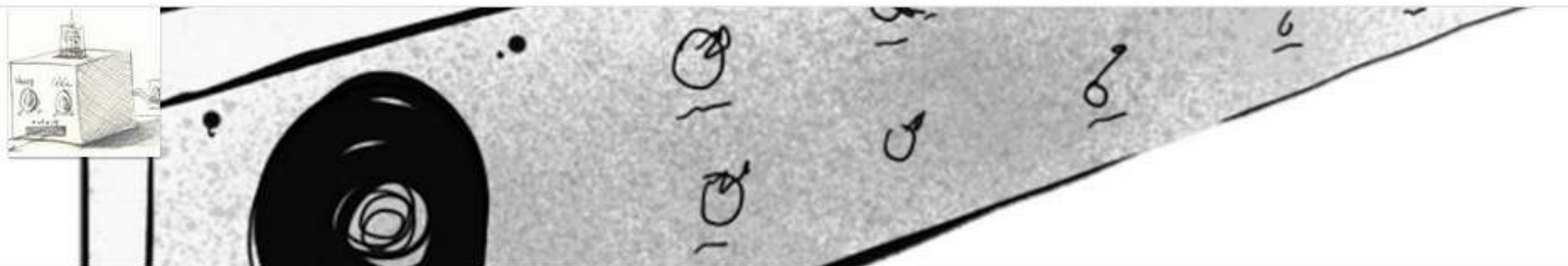


32 subscribers

6,520 views

Video Manager

https://www.youtube.com/channel/UC7wzmG64y2IbTUeWji\_qKhA



# The Bottomless Tune Box



Subscribe 32

The music on this channel arises from creative partnerships between musicians and a deep learning system that models music transcription ... Show more

## Uploads Public



Set #3 (fast reels)  
31 views • 1 week ago



Set #1 (jigs)  
69 views • 3 weeks ago



"Chicken Bits and Bits and Bobs"  
by Bob L. Sturm + folk-rnn  
68 views • 3 weeks ago



"Interlude" by Bob L. Sturm + folk-rnn  
26 views • 3 weeks ago



"The Humours of Time Pigeon" by Bob L. Sturm + folk-rnn  
41 views • 3 weeks ago

## Created playlists Public



C4DM concert (QMUL Nov. 18 2016)



Partnerships concert (May 23 2017)



Two short pieces and an Interlude in Concert



Bastard Tunes in Concert



C4DM concert (QMUL Nov 23 2015)



# An Unintentional Experiment!

Feedback Like 11.5M

Tuesday, Jun 20th 2017 8AM 69°F 11AM 75°F 5-Day Forecast

Daily **Mail**  
.com

Science & Tech

Home | U.K. | News | Sports | U.S. Showbiz | Australia | Femail | Health | **Science** | Money | Video | Travel | Columnists

Latest Headlines | Science | Pictures | Discounts

Login

## The future of music: 'Bot Dylan' AI writes its own catchy folk songs after studying 23,000 tunes

- Computer composes new tunes after being trained on 23,000 Irish folk songs
- This allowed AI to learn the patterns and structures that make for a catchy tune
- So far it has created over 100,000 new machine 'folk tunes', researchers say
- It marks a significant step forward for the capabilities of artificial intelligence



<https://www.dailymail.co.uk/sciencetech/article-4544400/Researchers-create-computer-writes-folk-music.html>



Duane\_1981, Preston, United Kingdom, 11 months ago

It's sounds very neat. It's missing the "human" element.



PaxRomana, Novi, 11 months ago

That's it?!!! I'm not impressed.



paevo, USA, United States, 11 months ago

Sounds like a robotic Irish jig....



Mikeyt1941, London, Canada, 11 months ago

Totally lifeless without warmth. Mind you much human tuneless junk that passes for music today isn't much better.

Click to rate



11



1



Fabrice, Manchester, United Kingdom, 11 months ago

No no no.



rocksnoop1, dover, United Kingdom, 11 months ago

Isn't music robotic enough these days?



pen, somewhere, United Kingdom, 11 months ago

Let's make all humans redundant, brilliant! Has everybody really lost their soul?!



Radar Also, Hemet, 11 months ago

This computerized "AI" is just so non musically untalented lazy nerds can infiltrate the world of true musicians who love, created, and write the music from the joy, hurt, and life emanating from their hearts.



Click to rate



7



1



# Evaluation: Unintentional experiments

Feedback Like 11.5M

Tuesday, Jun 20th 2017 8AM 69°F 11AM 75°F 5-Day Forecast

Daily Mail  
.com

Science & Tech

Home | U.K. | News | Sports | U.S. Showbiz | Australia | Femail | Her

Travel | Columnists

Latest Headlines | Science | Pictures | Discounts

Login

## The future of music: 'Bot D' writes its own catchy folk studying 23,000 tunes

- Computer composes new tunes after being trained
- This allowed AI to learn the patterns and structures that make
- So far it has created over 100,000 new machine 'folk tunes', researchers
- It marks a significant step forward for the capabilities of artificial intelligence

**BIAS**



# ***An Intentional Experiment***

How difficult will it be for a professional musician to produce an album using material generated by our system that will be judged successful within the idiom of Irish traditional music?

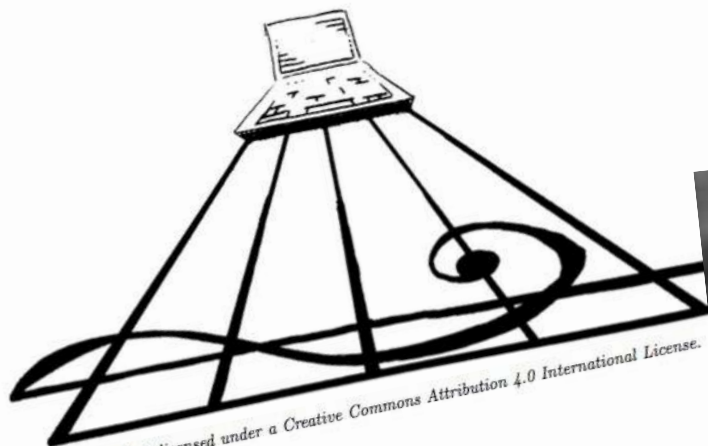


# 100,000 tunes in 34 volumes

The folk-rnn (v1) Session Book  
Volume 1 of 20\*

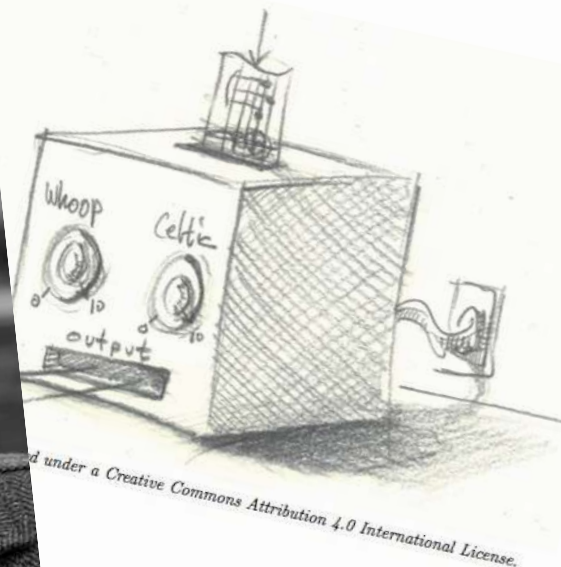


The folk-rnn (v3) Session Book  
Volume 1 of 4\*



This work is licensed under a Creative Commons Attribution 4.0 International License.

The folk-rnn (v2) Session Book  
Volume 1 of 10\*



This work is licensed under a Creative Commons Attribution 4.0 International License.



Daren Banarsë



**The reveal is  
coming soon...**





# “Let’s Have Another Gan Ainm”



Track listing:

1. Gan Ainm, Gan Ainm, Gan Ainm
2. The Drunken Landlady, Gan Ainm, Gan Ainm
3. Gan Ainm, Gan Ainm, Gan Ainm
4. Battle Of Aughrim, Gan Ainm, Lord Mayo
5. Gan Ainm, Gan Ainm, Tom Billy’s
6. Girls Of Banbridge, Gallowglass, Gan Ainm
7. The Blackbird, Gan Ainm, Mrs Galvin’s
8. Gan Ainm
9. Gan Ainm, Bunch of Green Rushes, Gan Ainm
10. Gan Ainm, Gan Ainm, Anthony Frowley’s
11. Gan Ainm, Toss the Feathers (II), Gan Ainm

<https://soundcloud.com/oconailfamilyandfriends>

Sturm, B. L. and Ben-Tal, O. (2018). *Let’s Have Another Gan Ainm: An experimental album of Irish traditional music and computer-generated tunes*. Technical report, KTH.



# “Let’s Have Another Gan Ainnm”

- 31 tunes in total arranged in 11 sets
- material of 20 tunes generated by *folk-rnn* models
- 11 tunes come from traditional repertoire
- Banarsë had access to all 100,000 tunes, but he took tunes from only six of the 34 volumes
- Banarsë and the musicians were free to do as they wished
- **More info:**
  - Ben-Tal, Harris, and Sturm, “How music Ai is useful: Engagements with composers, performers, and audiences”, *Leonardo*, 2020.

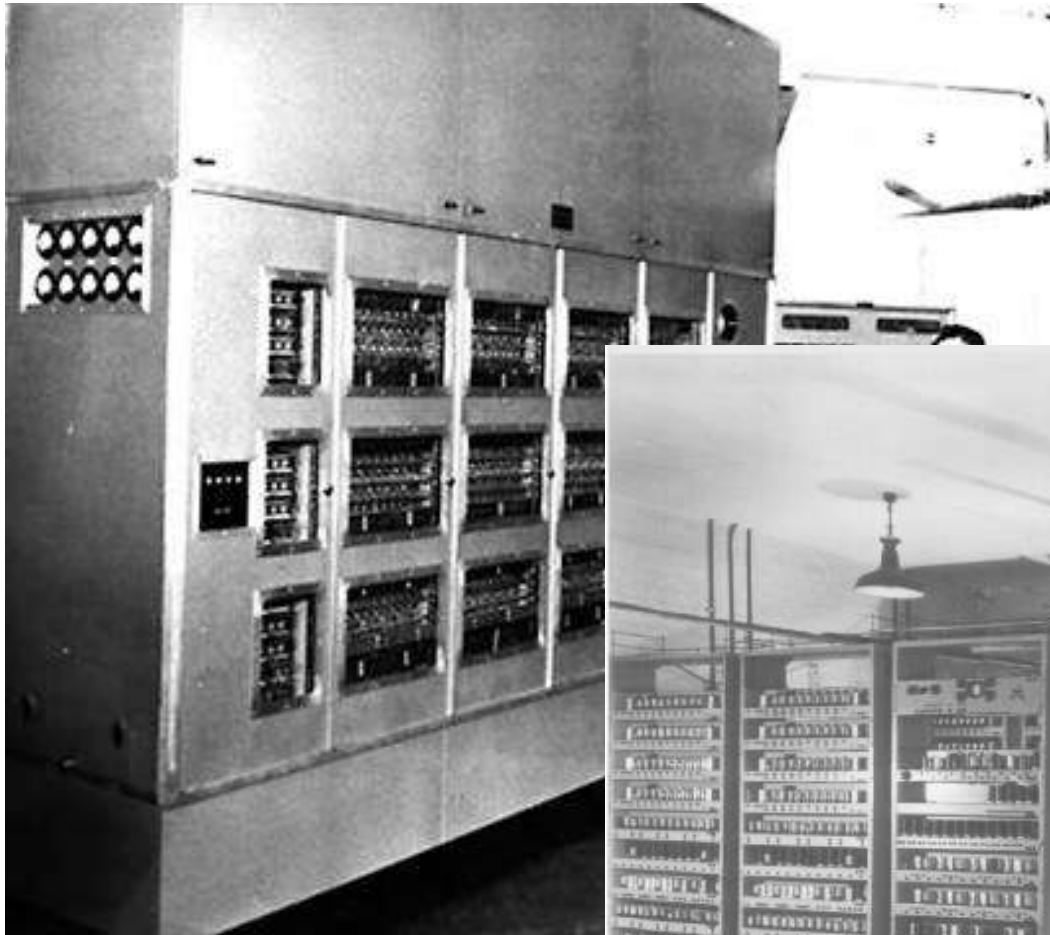
# Machines generating music is nothing new



*The Illiac Suite* (1957)

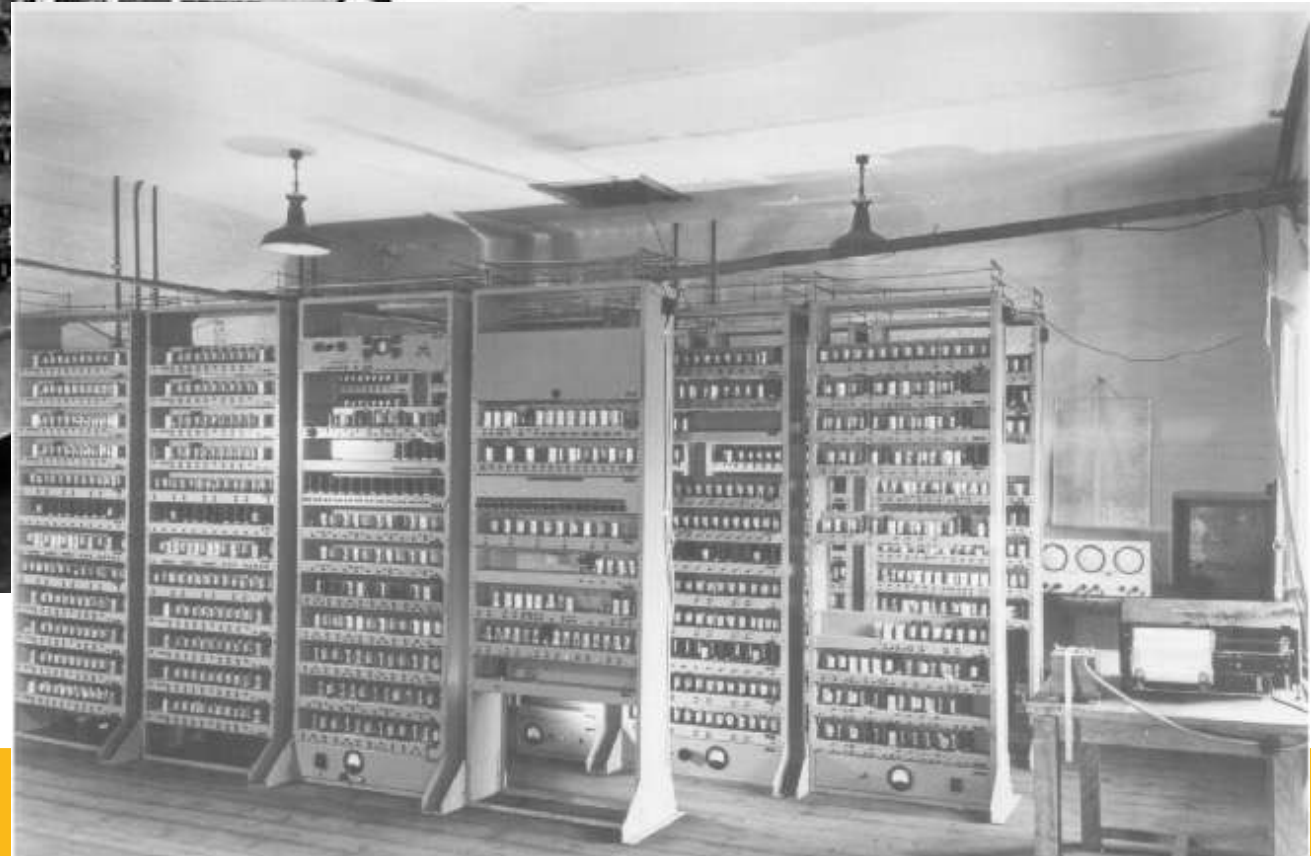


# Machines generating music is nothing new



*The Illiac Suite* (1957)

*Music from EDSAC* (1960)

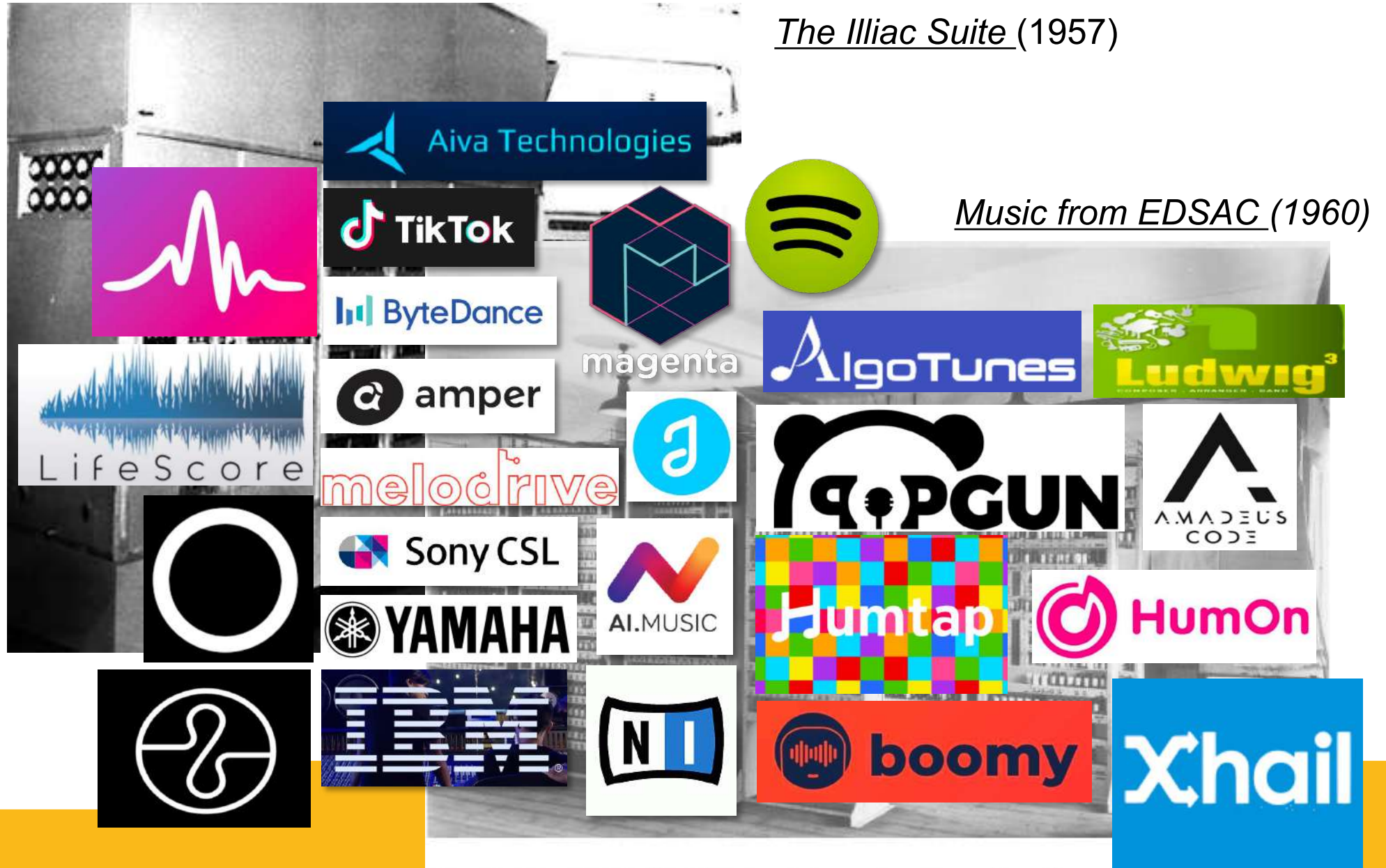


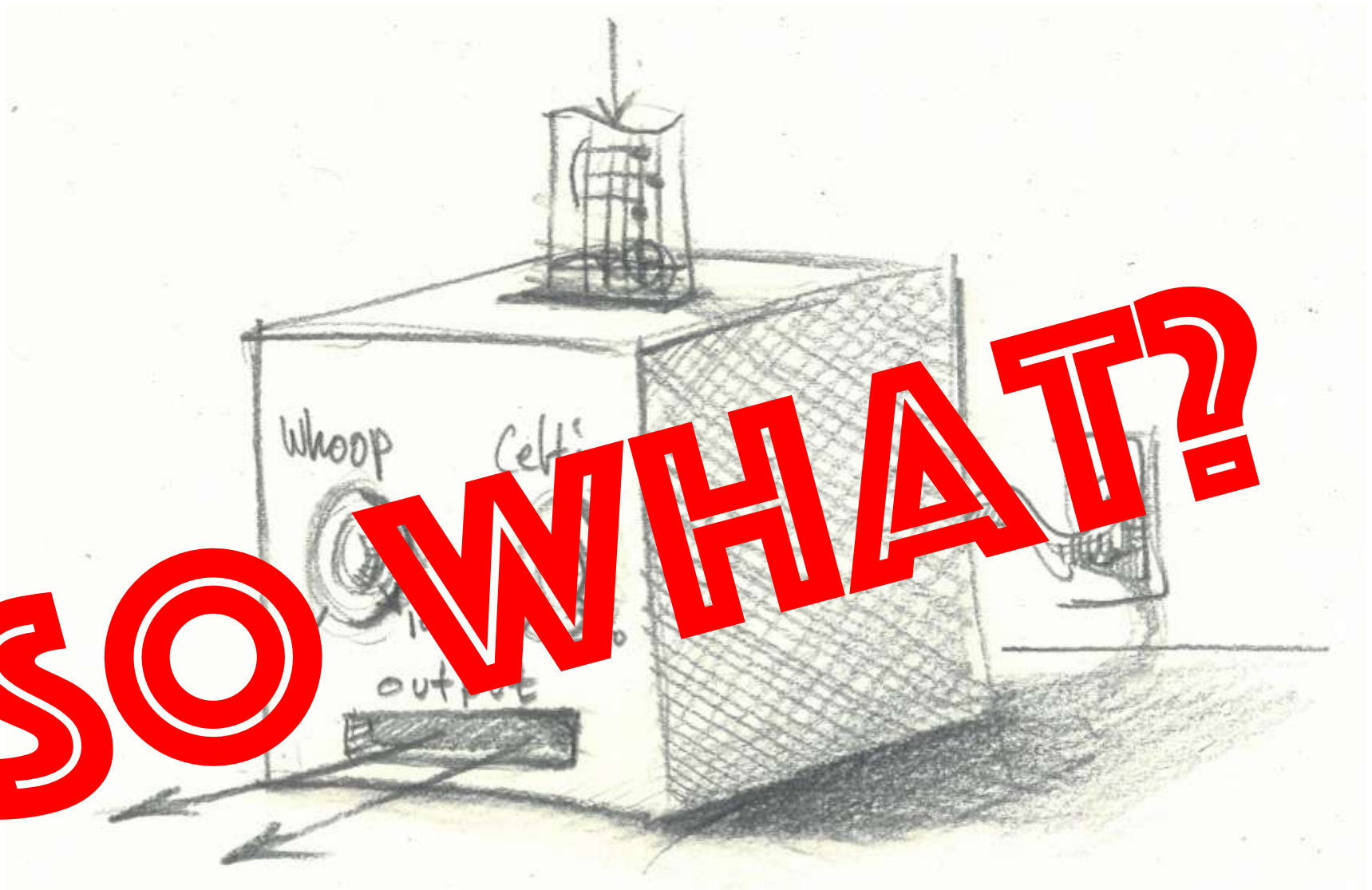


# Machines generating music is nothing new

The Illiac Suite (1957)

Music from EDSAC (1960)







# Let's get serious!

My Saturday morning fun developed into several interesting research questions:

- How far can we get with off-the-shelf (vanilla) ML?



**boblsturm** @boblsturm · 22h

Even more Endless Music Sessions

### Even more Endless Music Sessions

We have just injected 12,000+ new tunes into the Endless folk-rnn Traditional Music Session. Now you can audition sets assembled from tunes generated by thr...

[highnoongmt.wordpress.com](http://highnoongmt.wordpress.com)

← 1   ↻ 4   ❤️ 5   |||



**João Felipe** @seaandsailor · 18h

Paraphrasing @douglas\_eck, ABC is "cheat mode" for music generation :) folk-rnn doesn't have anything fancy besides a good representation.

← 2   ↻   ❤️ 1   ✉



# Let's get serious!

My Saturday morning fun developed into several interesting research questions:

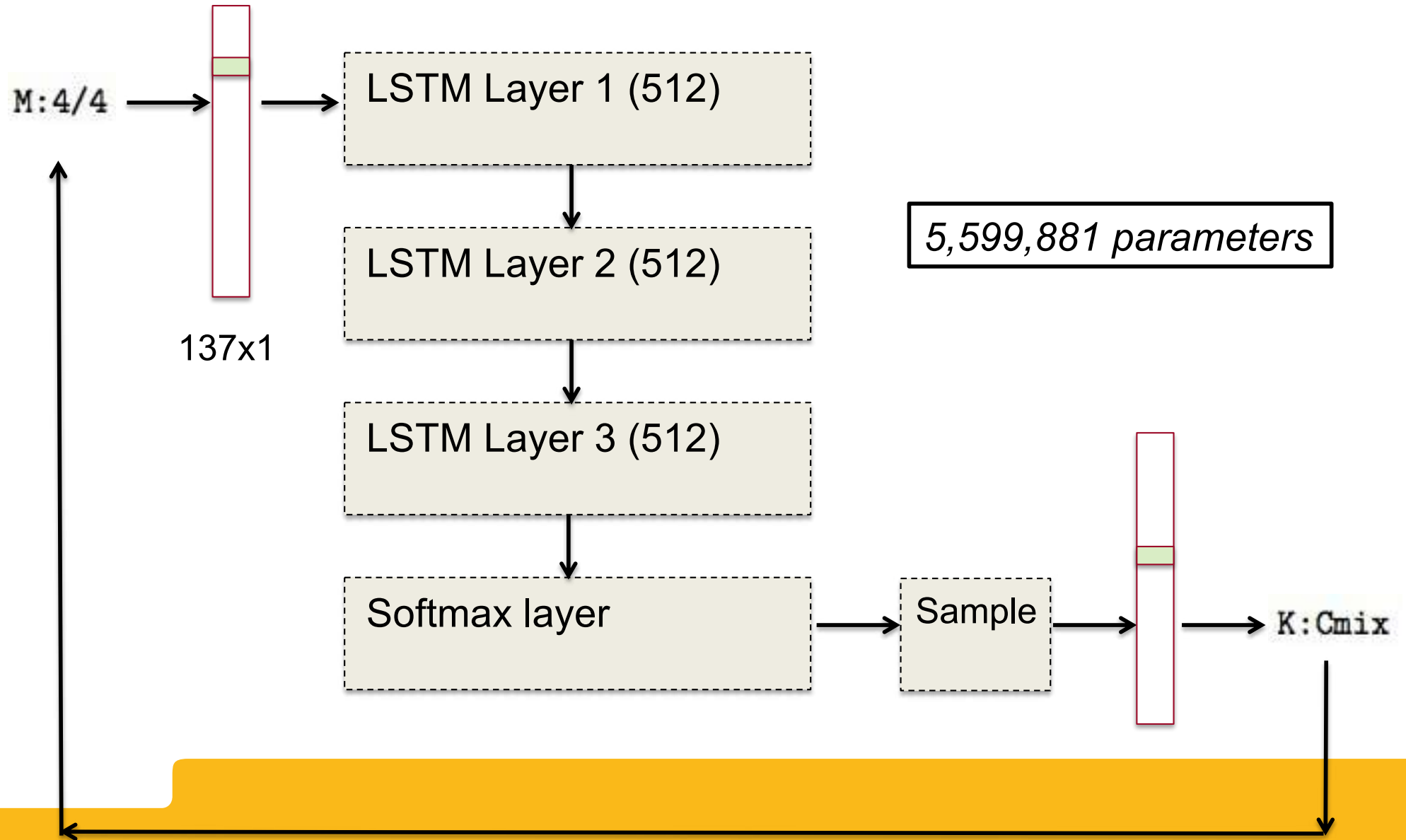
- How does the system work?
- Where is its knowledge in its 5.6 million parameters?

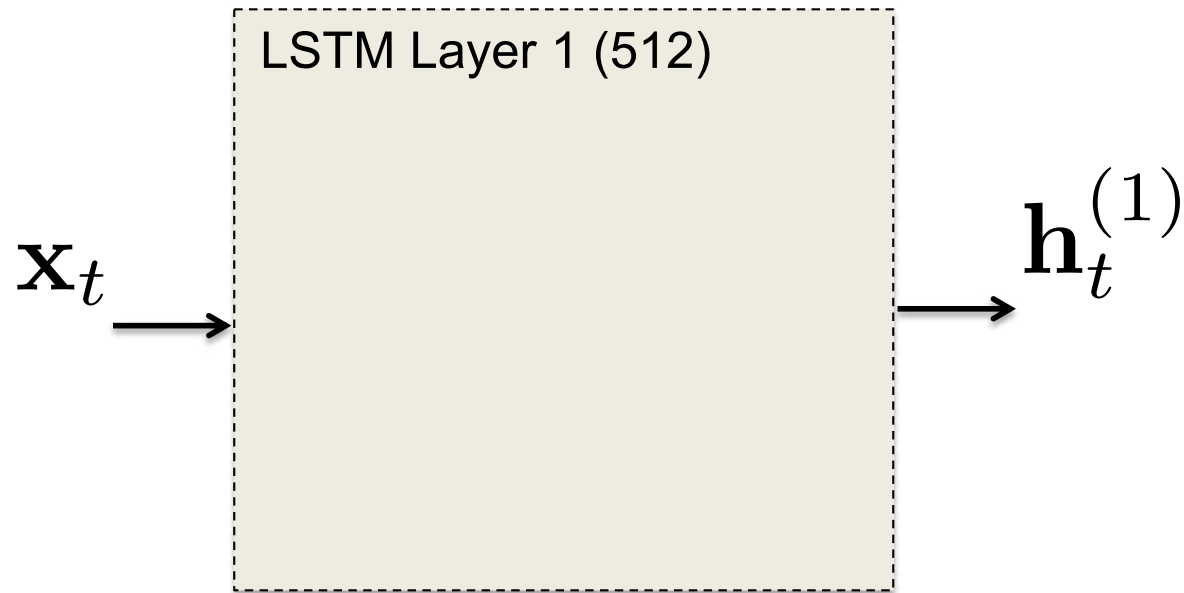
Sturm et al., “Machine learning research that matters for music creation: A case study,” *J. New Music Research* 48(1): 36–55, 2018.





# folk-rnn (v2): architecture





4 gates

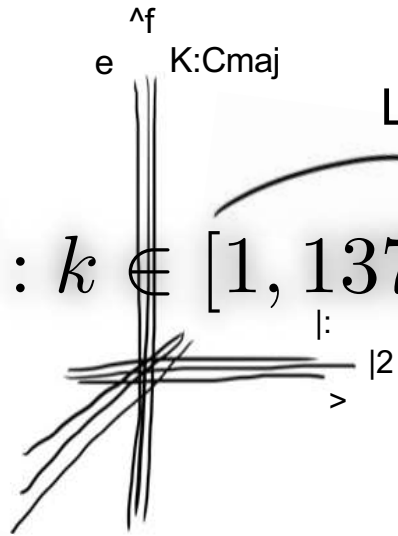
$$\left\{ \begin{array}{l} \mathbf{i}_t^{(1)} \leftarrow \sigma(\mathbf{W}_{xi}^{(1)} \mathbf{x}_t + \mathbf{W}_{hi}^{(1)} \mathbf{h}_{t-1}^{(1)} + \mathbf{b}_i^{(1)}) \\ \mathbf{f}_t^{(1)} \leftarrow \sigma(\mathbf{W}_{xf}^{(1)} \mathbf{x}_t + \mathbf{W}_{hf}^{(1)} \mathbf{h}_{t-1}^{(1)} + \mathbf{b}_f^{(1)}) \\ \mathbf{o}_t^{(1)} \leftarrow \sigma(\mathbf{W}_{xo}^{(1)} \mathbf{x}_t + \mathbf{W}_{ho}^{(1)} \mathbf{h}_{t-1}^{(1)} + \mathbf{b}_o^{(1)}) \\ \mathbf{c}_t^{(1)} \leftarrow \tanh(\mathbf{W}_{xc}^{(1)} \mathbf{x}_t + \mathbf{W}_{hc}^{(1)} \mathbf{h}_{t-1}^{(1)} + \mathbf{b}_c^{(1)}) \odot \mathbf{i}_t^{(1)} \\ \quad \quad \quad + \mathbf{f}_t^{(1)} \odot \mathbf{c}_{t-1}^{(1)} \end{array} \right.$$

hidden  
state

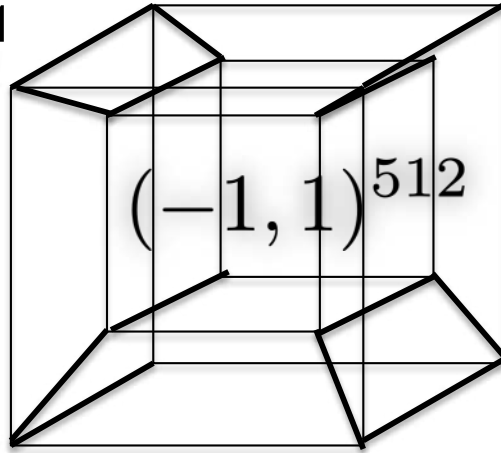
$$\mathbf{h}_t^{(1)} \leftarrow \tanh(\mathbf{c}_t^{(1)}) \odot \mathbf{o}_t^{(1)}$$



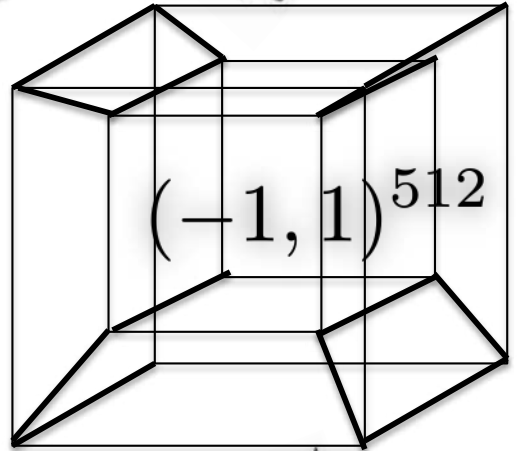
$$\{e_k \in \mathbb{R}^{137} : k \in [1, 137]\}$$



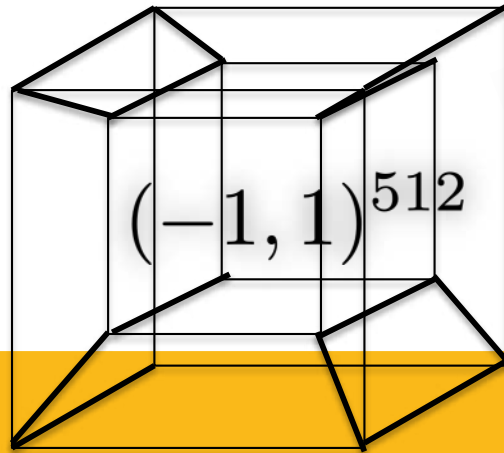
LSTM 1



LSTM 2

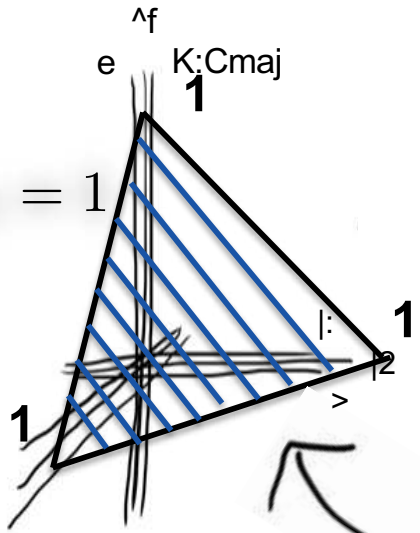


LSTM 3



sample

$$[0, 1]^{137} : \|\cdot\|_1 = 1$$



affine, softmax





# What do these 5,599,881 parameters mean? An analysis of a specific LSTM music transcription model, starting with the 70,281 parameters of its softmax layer

Bob L. Sturm, University of London UK

## How Stuff Works: LSTM Model of Folk Music Transcriptions

Bob L. Sturm<sup>1</sup>

A *folk-rnn* model is a long short-term memory (LSTM) that generates transcriptions. We evaluated these models using statistical analyses of their use in music practice and their behaviours precisely. This knowledge is essential for understanding them, and in this paper, we analyse the internal layer of a specific model. Key aspects of the model, for instance, that it is highly reliant on a specific layer to adjust the output and attenuate its probability.

The *folk-rnn* software (LSTM) network

We analyse the parameters of the first layer (L1) of an LSTM created by the *folk-rnn* software from crowd sourced music transcriptions (tunes) (Sturm et al., 2016).<sup>1</sup> The model has three hidden layers each of size 512 LSTM cells, and an input layer and softmax output layer of dimension 137. Each unit of the visible layers corresponds to one element of the vocabulary (token), e.g., M: 6/8 is 6/8 meter; K: Cmaj is C major mode; C is middle C; and | is a measure line. At step  $t$ , L1 transforms the input  $\mathbf{x}_t \in \{0, 1\}^{137} : \|\mathbf{x}_t\|_1 = 1\}$  into a hidden state  $\mathbf{h}_t^{(1)} \in (-1, 1)^{512}$  by (Graves, 2013):

$$\mathbf{i}_t^{(1)} \leftarrow \sigma(\mathbf{W}_{xi}^{(1)} \mathbf{x}_t + \mathbf{W}_{hi}^{(1)} \mathbf{h}_{t-1}^{(1)} + \mathbf{b}_i^{(1)}) \quad (1)$$

$$\mathbf{f}_t^{(1)} \leftarrow \sigma(\mathbf{W}_{xf}^{(1)} \mathbf{x}_t + \mathbf{W}_{hf}^{(1)} \mathbf{h}_{t-1}^{(1)} + \mathbf{b}_f^{(1)}) \quad (2)$$

$$\mathbf{o}_t^{(1)} \leftarrow \sigma(\mathbf{W}_{xo}^{(1)} \mathbf{x}_t + \mathbf{W}_{ho}^{(1)} \mathbf{h}_{t-1}^{(1)} + \mathbf{b}_o^{(1)}) \quad (3)$$

$$\mathbf{c}_t^{(1)} \leftarrow \tanh(\mathbf{W}_{xc}^{(1)} \mathbf{x}_t + \mathbf{W}_{hc}^{(1)} \mathbf{h}_{t-1}^{(1)} + \mathbf{b}_c^{(1)}) \odot \mathbf{i}_t^{(1)} + \mathbf{f}_t^{(1)} \odot \mathbf{c}_{t-1}^{(1)} \quad (4)$$

$$\mathbf{h}_t^{(1)} \leftarrow \tanh(\mathbf{c}_t^{(1)}) \odot \mathbf{o}_t^{(1)} \quad (5)$$

direction of the fifth rsv it amplifies the probability of the five measure tokens. We find that if we approximate  $\mathbf{W}_s$  by its first 30 singular vectors (reducing the subspace in which L3 operates from 137 to 30 dimensions) the model can still generate tunes with plausible local and global characteristics. If we only attenuate rsv 5 at L3 then the model is not able to correctly generate measure tokens. The model seems to rely on measure tokens to generate plausible tunes.

Since  $\mathbf{x}_t$  is one-hot corresponding to a token, we can relate each column of  $\mathbf{W}_{x*}^{(1)} \in \mathbb{R}^{512 \times 137}$  to one token, and groups of columns to token types. We want to know how these columns and the subspaces they span in  $\mathbb{R}^{512}$  relate to one another. Figure 1 shows the angles between all pairs of columns of each gate matrix. We see clear structures in those of the out and cell gates. Diagonals occur only for the pitch tokens, and relate tokens at octaves and enharmonics. The strongest off-diagonal line in (c) relates pitch tokens over three octaves without...



## Let's get serious!

My Saturday morning fun developed into several interesting research questions:

- How can we evaluate the resulting models?
- How can such models assist/hinder music creation?

*Involving practitioners, organizing workshops and concerts, composing ourselves and inviting others, cherry picking, ...*

*The models definitely assist me! But mileage may vary.*

Sturm et al., "Machine learning research that matters for music creation: A case study," *J. New Music Research* 48(1): 36–55, 2018.



# Let's get serious!

My Saturday morning fun developed into several interesting research questions:

- How can folk music improve ML research?
- What does this mean for traditional music?



# MUSAiC: 2020-2025 @ KTH

## *Music at the Frontiers of Artificial Creativity and Criticism*

Confronting questions and challenges at the frontier of the AI disruption/transformation of music

<https://musaiclab.wordpress.com>



European Research Council  
Established by the European Commission





# Ai Music Generation Challenge 2020

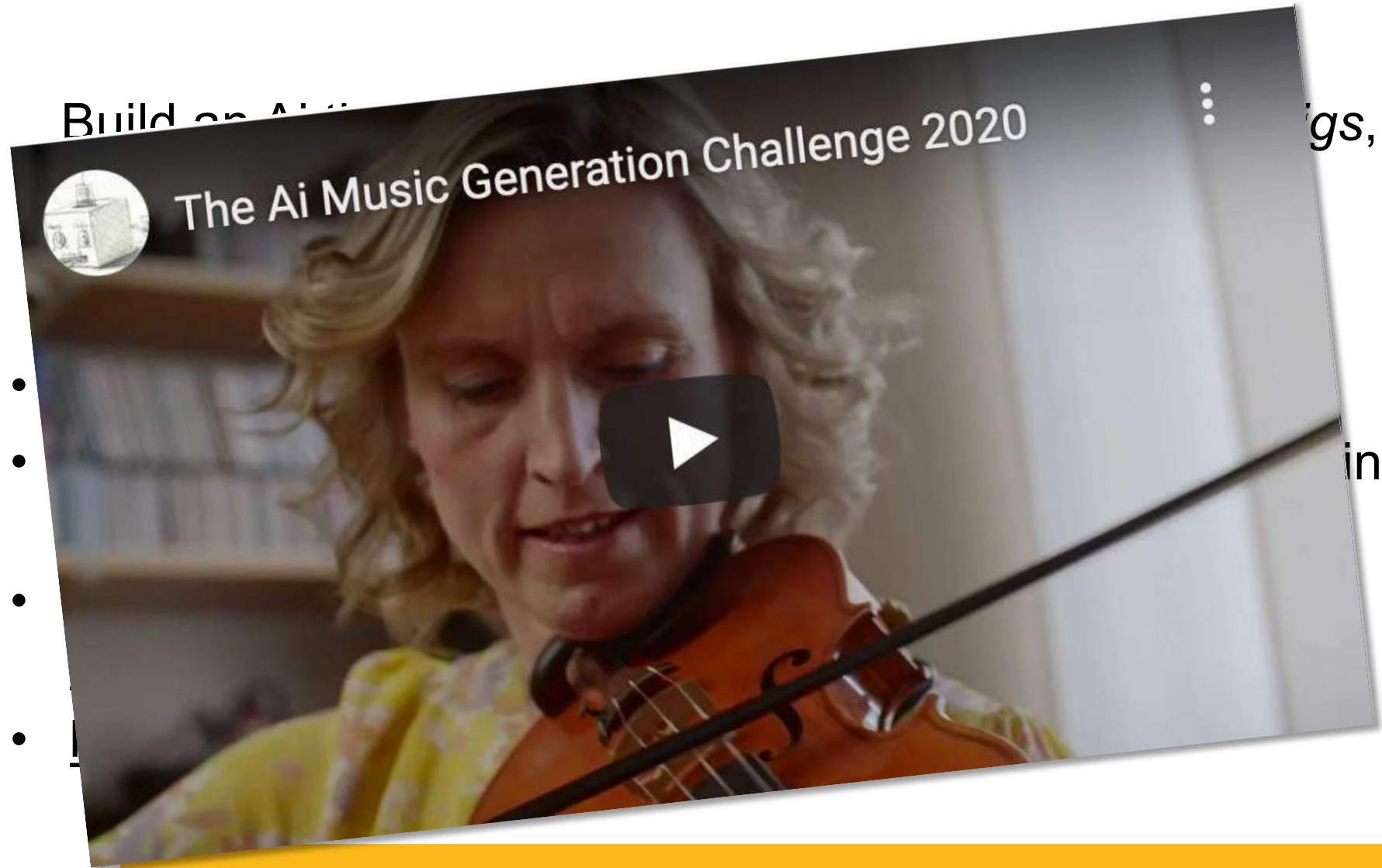
Build an Ai that generates the most plausible *double jigs*, as judged against the 365 published in O'Neill's *The Dance Music of Ireland* (1907).

- Up to two prizes awarded
- The panel of judges consisted of four (human) experts in Irish traditional music and performance
- A performance of the best AI jigs occurred at the 2020 Joint Conference on AI Music Creativity
- More information





# Ai Music Generation Challenge 2020



<https://youtu.be/KSoSyoEx6hc>



# Ai Music Generation Challenge 2020

Build an Ai that generates the most plausible *double jigs*, as judged against the 365 published in O'Neill's *The Dance Music of Ireland* (1907).

Three aims:

- to promote meaningful approaches to evaluating music Ai;
- to see how music Ai research can benefit from considering traditional music, and how traditional music can benefit from music Ai research;
- to facilitate discussions about the ethics of music Ai research applied to traditional music practices.



# Ai Music Generation Challenge 2021

Build an Ai that generates the most plausible ...

<https://github.com/boblsturm/aimusicgenerationchallenge2021>



# Ai Music Generation Challenge 2021

Build an Ai that generates the most plausible ...

*Slängpolskas!*



<https://github.com/boblsturm/aimusicgenerationchallenge2021>



# En själ-lös “svensk” låt!

## folkRNN

generate a folk tune with a recurrent neural network

PRESS TO GENERATE TUNE

Compose

MODEL

folkwiki.se

TEMPERATURE

SEED

1



352794

METER

MODE

3/4

D Minor

INITIAL ABC

Enter start of tune in ABC notation

Bm Em A D

D A F# Em Bm

D A D F#

G a Bm Em A Bm

<https://themachinefolksession.org/tune/551>



## Tunes from the Ai Frontiers

Most recent tunes

### Week 26: The dog ate a raisin so call the vet (folk-rnn v2 + Sturm)

[A dog and a raisin] is better than a dog and a piece of cardboard. Information courtesy of GPT-2 True story: Carla got herself a nice big bowl of muesli and vanilla yogurt and then went to get a cup of tea. She came back to the table to find our dog Shoogee with her

Continue reading →

March 24, 2021

### Week 25: Shoogee Take Another Shoe (folk-rnn v2 + Sturm)

[Dogs stealing shoes] may appear on the street as a casual problem, but it is extremely dangerous. They have strong jaws

About Videos Praise FAQ



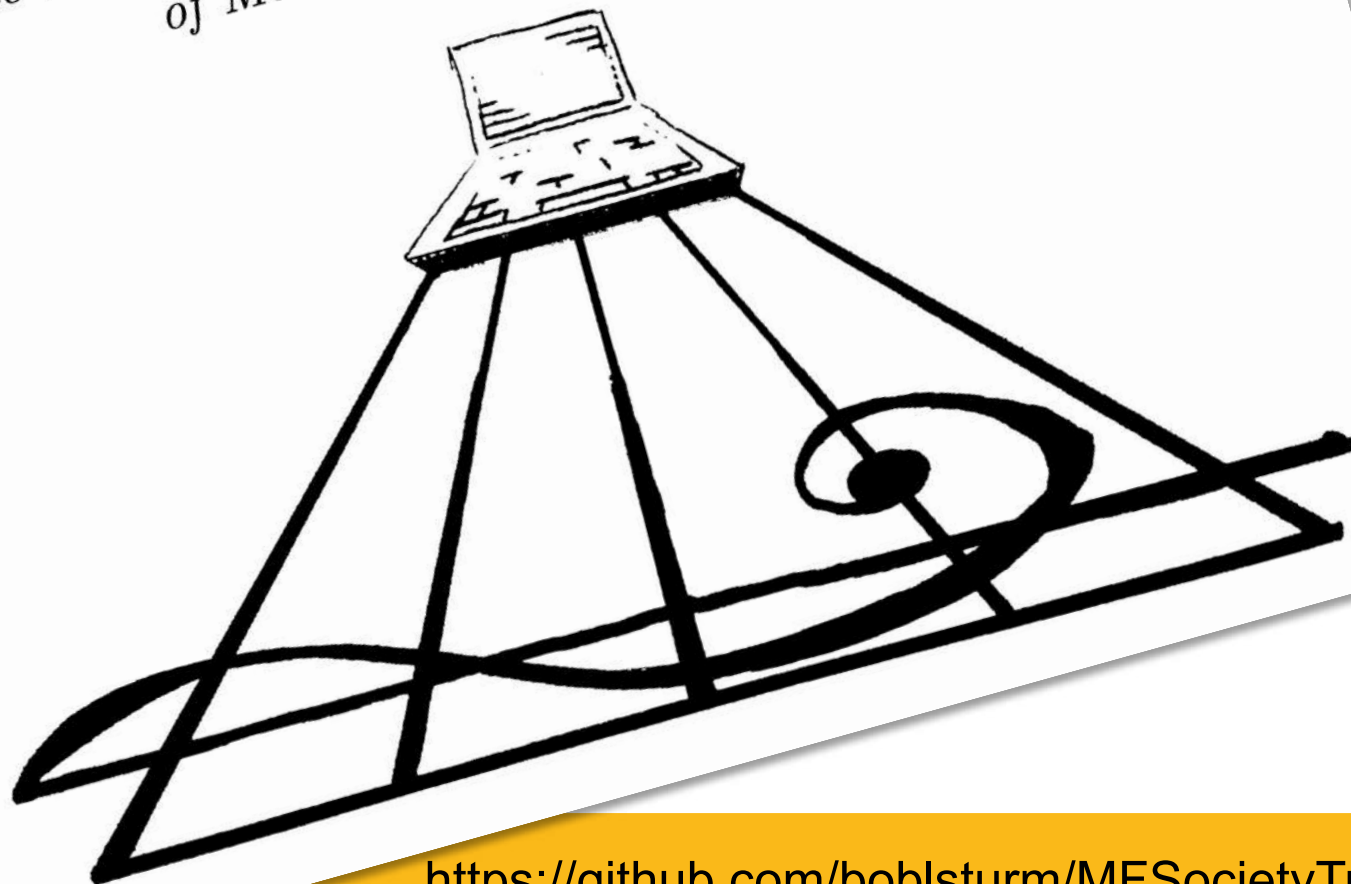
### About Me

Hi. I'm Bob, the unelected President of The Society for the Preservation and Promotion of Machine Folk Music (v1.1). This blog documents my journey learning folk music generated by Ai and promoting it. This "poetic research" is supported by ERC-2019-COG No. 864189 MUSAIC: Music at the Frontiers of



# Want to do more?

*The Society for the Preservation and Promotion  
of Machine Folk Music (v1.1)*



<https://github.com/boblsturm/MFSocietyTunebook>

<https://www.facebook.com/groups/TheMFSociety>