

Machine Learning
@devfest

```
book-nav-toggle"  
idden="" fixed="" aria-label='Hide  
Hide side navigation"
```

The Age of Digital Da Vinci: All about of Image Generation



Abhik Sarkar
Machine Learning, Cloudastructure

Agenda for today

1. Prompt Engineering
2. Advanced Image Prompting
3. AutoEncoders
4. GANs
5. Diffusion Models
6. Very Advanced Image Prompting
7. Future & Beyond



Midjourney



Stable Diffusion

Machine Learning
@devfest

Prompt Engineering



```
on class= 'devsite-book-nav-toggle'
haspopup='menu' hidden=''' fixed=''' aria-label='Hide
navigation' data-title='Hide side navigation'
-expanded='true'><span class='material-icons
' /></span></button>
```



```
on class= 'devsite-book-nav-toggle'
-haspopup='menu' hidden=''' fixed=''' aria-label='Hide
navigation' data-title='Hide side navigation'
-expanded='true'><span class='material-icons
' /></span></button>
```

What Is Prompt Engineering?

A prompt is the input you provide, typically text, when interfacing with an AI model like ChatGPT or Midjourney.

Prompt engineering is the process of discovering prompts which reliably yield useful or desired results.



Prompt Engineering - 101

- Give Direction: Describe what you're imagining, to get an output matching your vision.
- Specify Format: Define the response you want, and minimize time spent parsing errors.
- Provide Examples: Insert examples in your prompts, to improve the reliability of the output.
- Divide Labor: Split tasks into multiple prompts, chained together for complex goals.

Give Direction

Can you guess prompt

For this image?



Give Direction

people in a business
meeting



Give Direction

people in a business
meeting



Give Direction

stock photo of
business meeting of
**four people gathered
around a campfire
outdoors in the
woods, Panasonic,
DC-GH5**



Give Direction

people in a business
meeting, by **greg
rutkowski**



Specify Format

business meeting of
four people watching
on MacBook on top of
table, **in minecraft**



Provide Examples

stock photo of
business meeting of
four people watching
on white MacBook on
top of glass-top
table, Panasonic,
DC-GH5



Result



Divide Labor

Michael is at that really famous museum in France looking at its most famous painting. However, the artist who made this painting just makes Michael think of his favorite cartoon character from his childhood. What was the country of origin of the thing that the cartoon character usually holds in his hand?

class="material-icons" data-kind="parent" data-rs="4">menu

aspopup="menu" hidden="" fixed="true" data-bbox="0 0 340 130" data-label="Text">

avigation" data-title="Hide side navigation" data-bbox="0 0 340 130" data-label="Text">

xpanded="true">

/button>

JAPAN!

Divide Labor

- Michael is at that really famous museum in France looking at its most famous painting.

Mona Lisa

Divide Labor

- the artist who made this painting

Leonardo Da Vinci

Divide Labor

- Michael think of his favorite cartoon character from his childhood

Which can is related to Leonardo Da Vinci

Leonardo from Teenage Mutant Ninja Turtle

Divide Labor

- What does Leonardo from Teenage Mutant Ninja Turtle hold in his Hand?

Katana

Divide Labor

- What is country of origin of Katana

JAPAN!

`"class="time talk-ended single-
" class="talk-name">...
" class="description">...`

Advanced Image Prompting



Machine Learning @devfest

```
"class="time talk-ended single"  
"class="talk-name">...  
"class="description">...
```



AIM: To write advanced Prompt like this

a colorful digital artwork of a woman's head and branches, in the style of haunting houses, graceful surrealism, depictions of urban life, dark sky-blue and orange, portraiture with hidden meanings, caricature-like illustrations, metropolis meets nature --ar 69:128 --s 750 --niji 5



Specify Format Advanced

an oil painting
of a business
meeting



Specify Format Advanced

an oil painting
of a business
meeting, textured
oil-on-canvas
using thick
impasto and
swirling dynamic
brushstrokes



Specify Format Advanced

an ancient
egyptian
hieroglyph of a
business meeting



```
on class= 'devsite-book-nav-toggle'
-haspopup='menu' hidden=''' fixed=''' aria-label='Hide
navigation' data-title='Hide side navigation'
-expanded='true'><span class='material-icons
' /></span></button>
```

Image Specific Prompt Design

Art Style Modifiers

Negative Prompts

Weighted Terms

Inpainting



Art Style Modifiers

illustration of a
dragon, in the
style of greg
rutkowski



Negative Prompts

allow you to
specify what you
don't want in the
image.

oil painting in the
style of rembrandt



oil painting in
the style of rembrandt
--no frame, wall



Weighted Terms

dali :: 0
van gogh :: 1



dali :: 0.2
van gogh :: 0.8



dali :: 0.4
van gogh :: 0.6



dali :: 0.6
van gogh :: 0.4



dali :: 0.8
van gogh :: 0.2

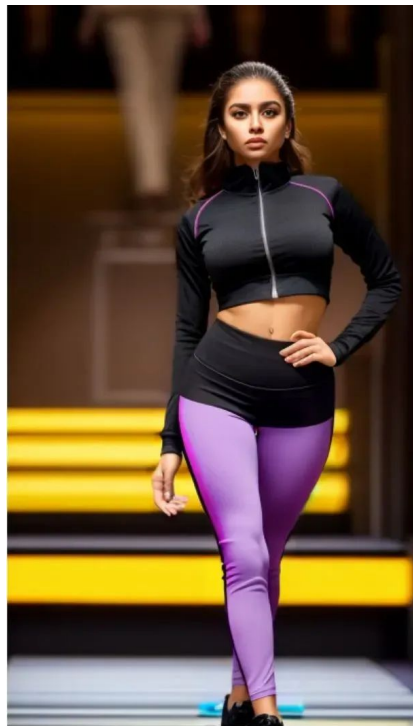
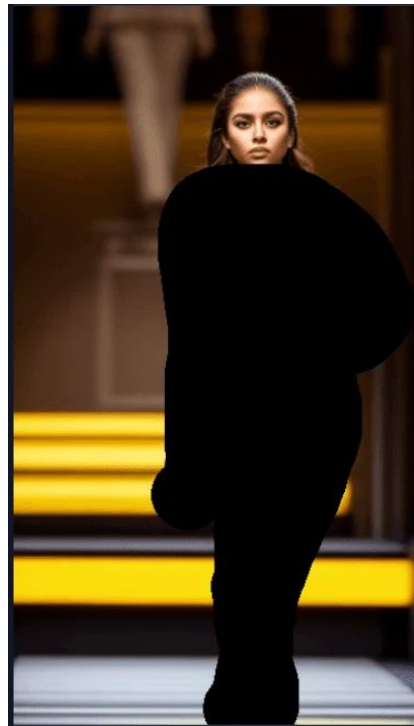
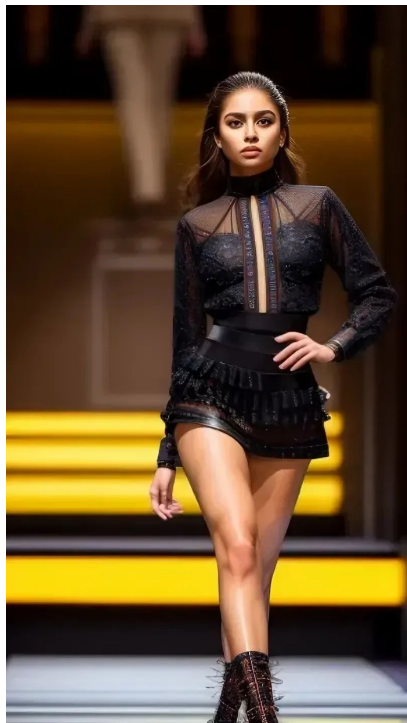


dali :: 1
van gogh :: 0



Inpainting

sportive
style, sport
outfit, sport
wear, sport
clothes,
(black yoga
pants:1.3),
sneakers

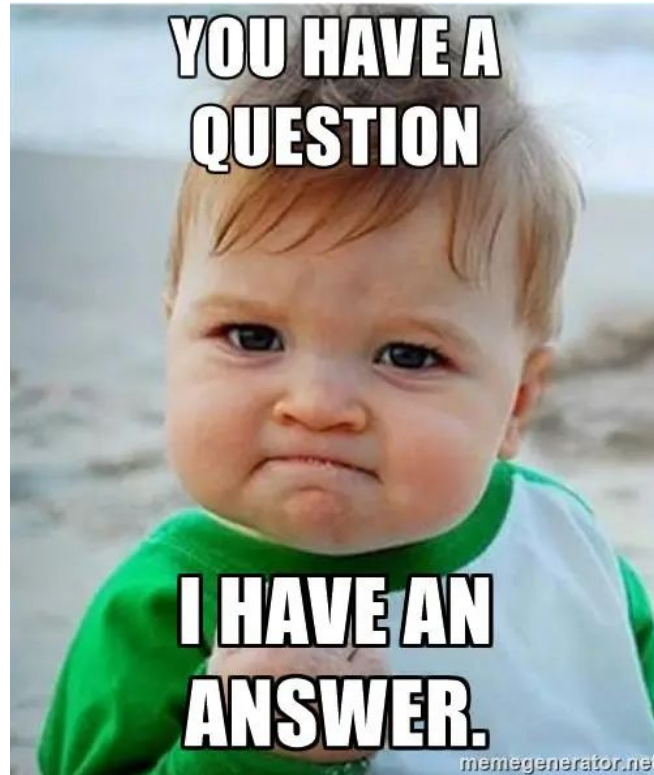


```
on class= 'devsite-book-nav-toggle'  
-haspopup='menu' hidden='' fixed='' aria-label='Hide  
navigation' data-title='Hide side navigation'  
-expanded='true'><span class='material-icons  
' /></span></button>
```

HOLY GRAIL



```
on class= 'devsite-book-nav-toggle'
-haspopup='menu' hidden='' fixed='' aria-label='Hide
navigation' data-title='Hide side navigation'
-expanded='true'><span class='material-icons
' /></span></button>
```

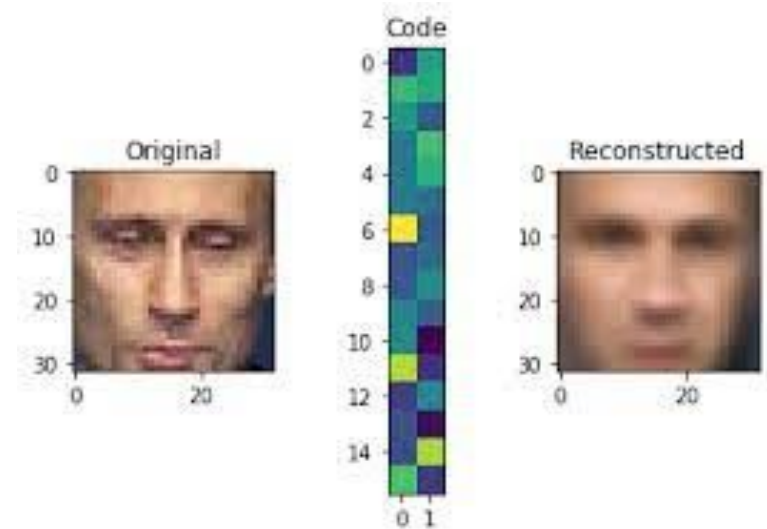
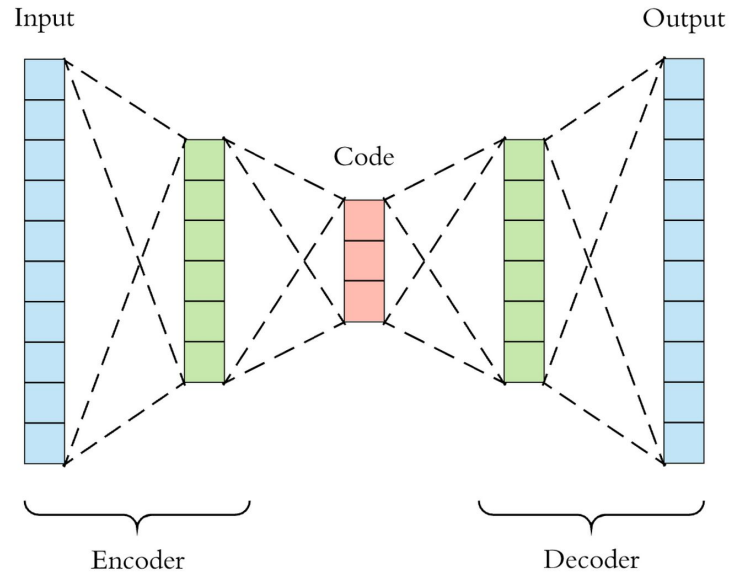


"class='time talk-ended single-
"class='talk-name'>...
"class='description'>..."

Auto Encoders



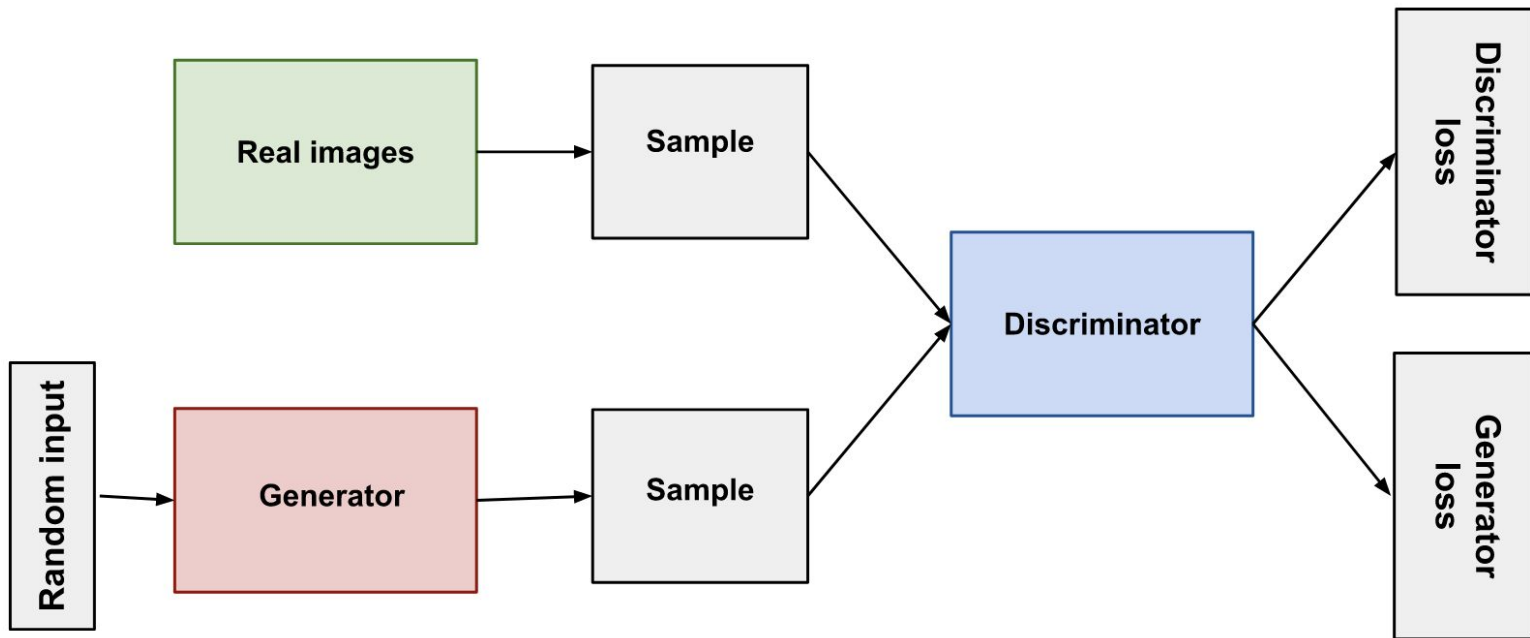
Autoencoder Compression



Generative Adversarial Network (GANs)



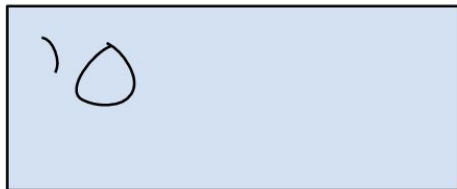
GAN Architecture



Generated Data

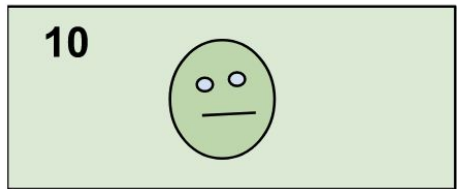
Discriminator

Real Data



FAKE

REAL



FAKE

REAL



REAL

REAL



"class="time talk-ended single-
"class="talk-name">...
"class="description">...

Diffusion Models



Diffusion Models

Denoising diffusion models

- **Forward / noising process**

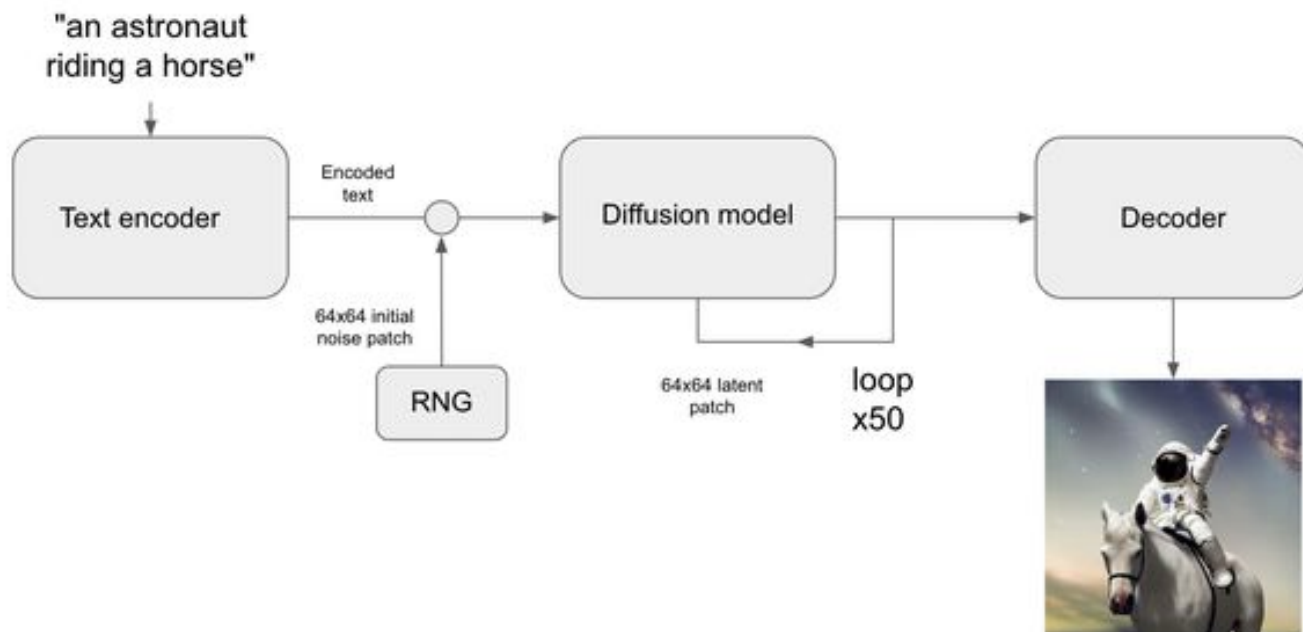
- Sample data $p(x_0) \rightarrow$ turn to noise



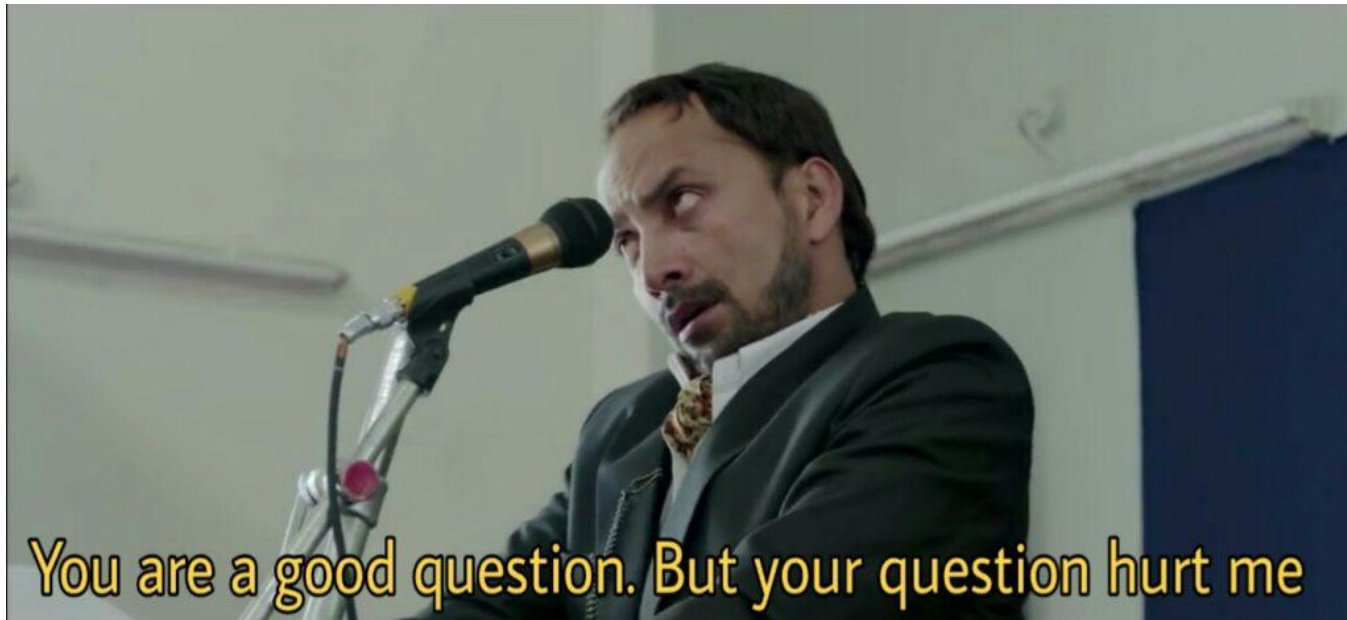
- **Reverse / denoising process**

- Sample noise $p_T(x_T) \rightarrow$ turn into data

Diffusion Models



```
on class= 'devsite-book-nav-toggle'
-haspopup='menu' hidden=''' fixed=''' aria-label='Hide
navigation' data-title='Hide side navigation'
-expanded='true'><span class='material-icons
' /></span></button>
```



You are a good question. But your question hurt me

Very Advanced Image Prompting




```
on class= 'devsite-book-nav-toggle'  
-haspopup='menu' hidden='' fixed='' aria-label='Hide  
navigation' data-title='Hide side navigation'  
-expanded='true'><span class='material-icons  
' /></span></button>
```

Advanced Techniques

- ControlNet
- LoRA



ControlNet

- Takes overall structure from one image.
- ControlNet takes in the text prompt and transforms it into a set of guidance signals that influence how the art is composed.

Input



Intermediate

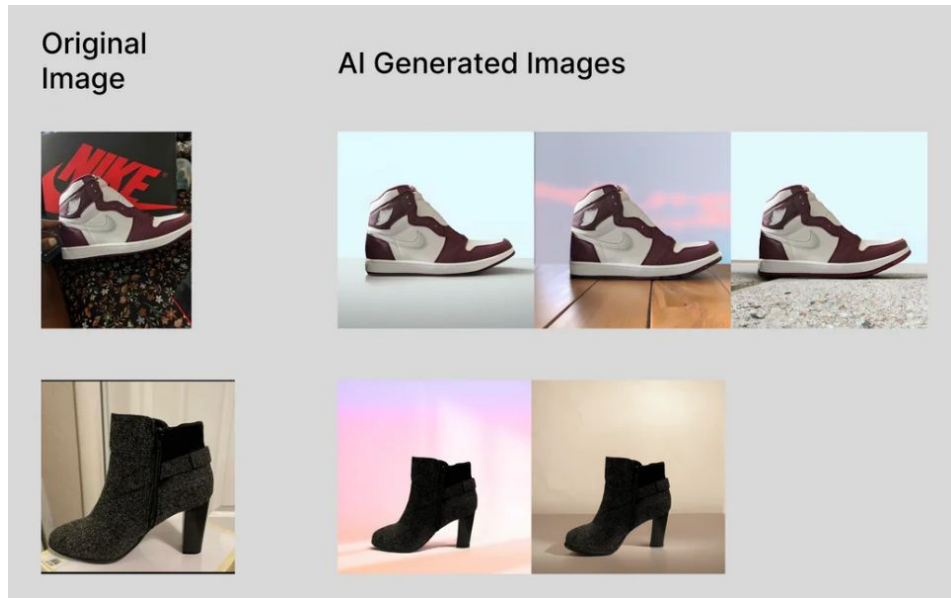


Output



LoRA

- Build small plug-in modules that teach the model to focus on new concepts and styles
- Specific type of output.



Machine Learning
@devfest

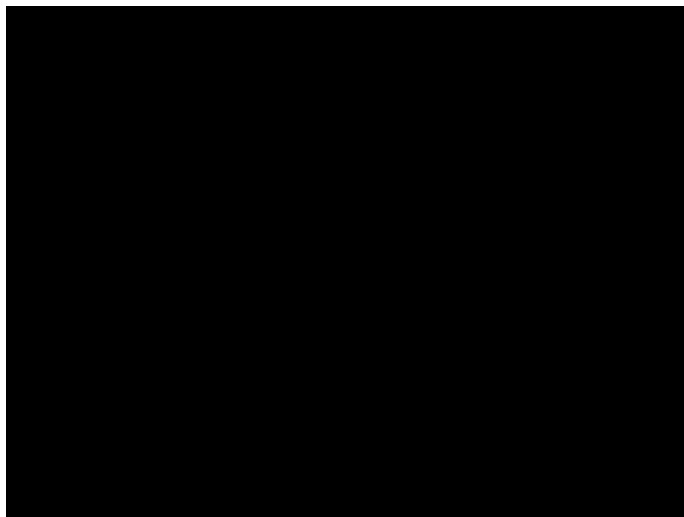
`"class="time talk-ended single-
" class="talk-name">...
" class="description">...`

Future & Beyond



Machine Learning @devfest

```
"class="time talk-ended single-  
"class="talk-name">...  
"class="description">...
```



Machine Learning
@devfest

`"class="time talk-ended single-
" class="talk-name">...
" class="description">...`

video



Machine Learning
@devfest

```
"class="time talk-ended single"  
"class="talk-name">...  
"class="description">...
```

End of Presentation



ANY QUESTIONS?

memecreator.org

4/7

Machine Learning
@devfest

`"class="time talk-ended single-
" class="talk-name">...
" class="description">...`



That's all ~