

When AI Gets Creative: Controlling Quality in Mass-Generated Product Images

PD Dr-Ing. habil. Jürgen Nützel, CEO, jn@deenovum.com

Fahad Mansuri, fm@deenovum.com

APPIS 2026
8th International
Conference
Las Palmas de Gran
Canaria, Spain
January 19-23, 2026



Find me on LinkedIn

Abstract

LLM-powered image generation promises a revolution for B2B commerce: faster product onboarding, automated visual enhancements, and scalable content production. The reality, however, is messy. Even with consistent prompts, image outputs fluctuate—sometimes subtly, sometimes dramatically. Shadows shift, colors drift, proportions distort, and occasionally the product... isn't the product anymore.

This session confronts the core problem of instability in AI-generated imagery and introduces a new automated benchmarking framework for detecting and quantifying quality issues at scale. You'll see how objective scoring, consistency checks, and visual metrics can transform unpredictable AI creativity into a controlled, reliable production process—without sacrificing speed or automation.

Growing Digital, *Together.*

- Digitalization pioneers since 2000
- Spin-off of Fraunhofer IDMT & TU Ilmenau
- CEO, PD and main owner: Dr. Jürgen Nützel
- E-Commerce Experts (> 20 years)
- Intershop Partner (for more than 10 years)
- Cloud Computing (AWS since 2013)
- AWS (Amazon Web Services) Partner since 2017
- Odoo Silver Partner (two years)
- 15 employees in Germany (deenovum Technologies AG)
- 8 employees in Gujarat India (deenovum Technologies India Pvt. Ltd.)



A look at deenovum India



In October I was in India to inaugurate our new office there.
It took me two tries to crack the coconut.
A German might need to get used to some of the symbols again.





*The idea to use generative AI came
from a customer in the electrical
wholesale trade.*



The hometown of LIMMERT is Salzburg



Their new Autostore System



In 2023, I presented an innovative AI-based search at Limmert.
This sparked interest in more AI solutions ...



Managing Director Thomas Canek
Owner Birgit Limmert
Father Karl Limmert

With the customer Gbr. Limmert AG to the solution...



The customer

The electrical wholesaler
Gebrüder LIMMERT in
Salzburg, would like us to
revise its product data for
more than one million
products.



Our idea

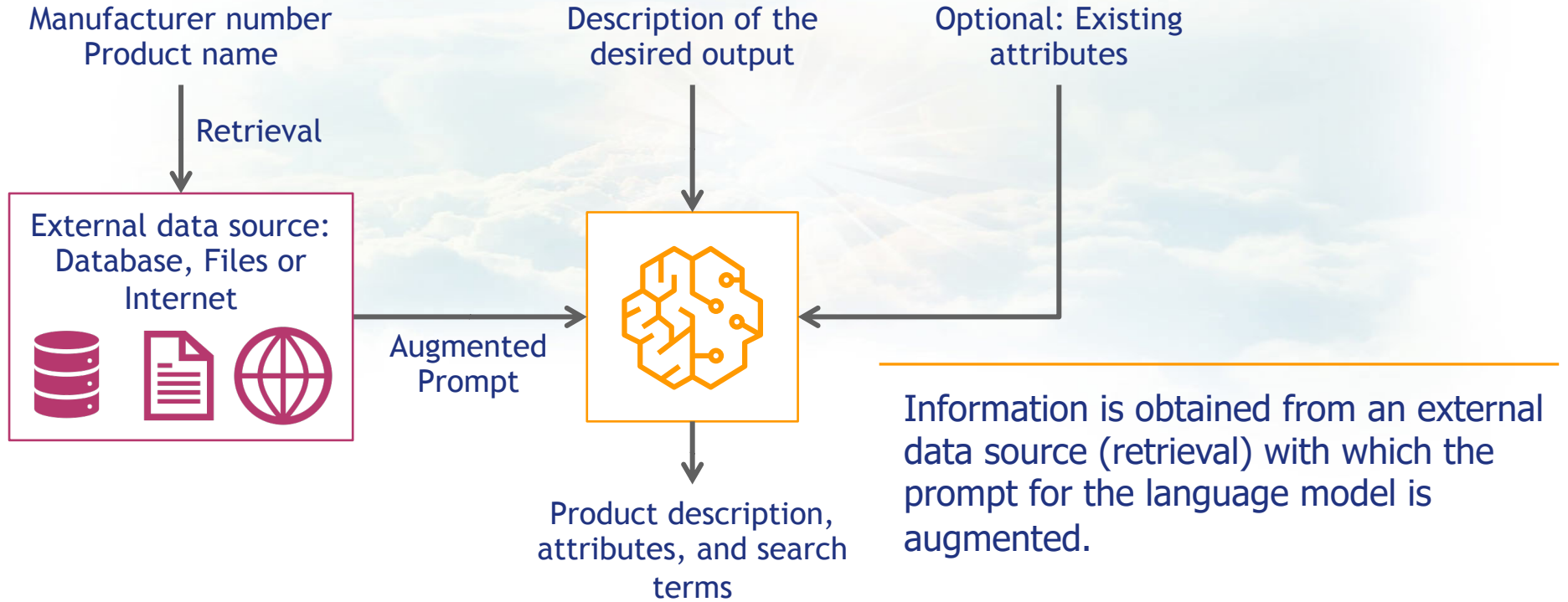
of using generative AI models
to create product data
directly seemed to work at
first glance.



Our solution

has become a lot more
complex in order to meet the
requirements of electrical
experts.

Retrieval-Augmented Generation (RAG)



*A look inside their online shop,
which we built.*

1384 COROPLAST ORANGE 0,10X15X10 TYP 301



Artikelnummer : 01046949
Herkunftsland : DE
Zolltarifnr : 39191012
EAN: 4021393002415
Meine Artikel Nr.:

COROPLAST

Mögliche Verpackungseinheiten

1 Schachtel (20 Roll)
(0.05 m X 0.1 m X 0.186 m)
Gewicht: 0.5 kg

LimCoin Warenkorb Aktion



Warenkorbwert
ab 100 €
ab 250 €
ab 500 €

erhältst du
6 LimCoins
17 LimCoins
40 LimCoins

pro 100 Roll
124,29 €
exkl. USt.

Farbe

orange

Roll

-

20

+

Warenkorb

Meine Bestellungen

In Bestellvorlage aufnehmen

Sie hätten nach diesem Produkt anders
gesucht?

AI-generated data

LIMMERT-KI ✨ Beschreibung

Das Coroplast Isolierband Typ 301 in orange bietet eine zuverlässige Lösung für vielfältige Anwendungen. Mit einer Breite von 15 mm und einer Länge von 10 m ist es ideal für das Abdecken und Isolieren von Drähten und Kabeln geeignet. Hergestellt aus hochwertigem Polyvinylchlorid (PVC), ist dieses Band temperaturbeständig bis 105 °C und selbstverlöschend, was es zu einer sicheren Wahl für den Einsatz in verschiedenen Umgebungen macht. Obwohl es nicht isolierend oder UV-beständig ist, überzeugt es durch seine Robustheit und Vielseitigkeit. Perfekt für Heimwerkerprojekte, Reparaturen und mehr, ist das Coroplast Isolierband ein unverzichtbares Werkzeug in jedem Haushalt oder Werkstatt.

Technische Merkmale (Limmert KI extrahiert):

Länge:	10 m	Temperaturbeständig bis:	105 °C
Breite:	15 mm	Stärke:	0,1 mm
Isolierend:	Nein	Selbstverschweißend/Amalgamierend:	Nein
Werkstoff:	Polyvinylchlorid (PVC)	UV-beständig:	Nein
Selbstverlöschend:	Ja	Farbe:	orange

Produktmerkmale

Previously
existing data

Werkstoff:	sonstige	Breite:	15 mm
Länge:	10 m	Farbe:	orange
Isolierend:	Nein	UV-beständig:	Nein
Temperaturbeständig bis:	105 °C	Stärke:	0.1 mm
Selbstverschweißend/Amalgamierend:	Nein	Selbstverlöschend:	Ja
Doppelseitig:	Nein	Geeignet für Hochspannung:	Ja
Vulkanisierend:	Nein		



And what about the product images?

Since the lamp is to be sold in Salzburg, place the lamp in front of a suitable landscape.

Nachgedacht für 8 Sekunden >

Bild erstellt



Add a hat typical of Salzburg to the lamp.

Bild erstellt



Hier ist deine Philips-Lampe mit einem typischen Salzburger (Tyrolean) Hut! Wenn du noch Details anpassen möchtest – z.B. andere Farbe oder Federform – sag einfach Bescheid. ↓

Some Real Use-cases in Industrial Applications

1

Extend Image

- Add shop logo
- Add watermark
- Add or remove background

2

Modify Product Image

- Improve image quality
- Change attributes like
 - color
 - length
 - number of wires
 - ...

3

Marketing

- Product in usage scenario
- Combine with persons

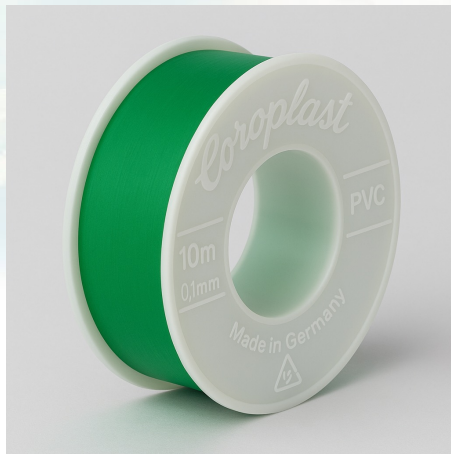
Example scenarios using orange tape

PROMPT:

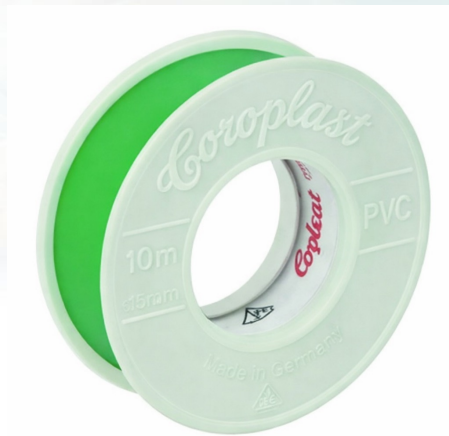
Change the colour to green



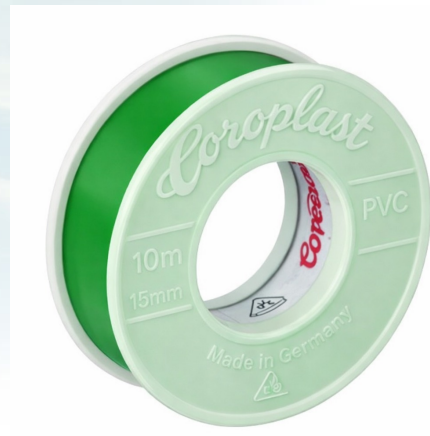
Original image



gpt-image-1
(old model)



gpt-image-1.5



Chatgpt-image-latest
(same model used in
ChatGPT)

Marketing scenarios

PROMPT:

Change the colour to green and put it in the hand of a model



Original image



gpt-image-1.5



Chatgpt-image-latest

Challenges and problems we have to deal with

1

Quality

LLMs cannot guarantee quality and reproducibility

2

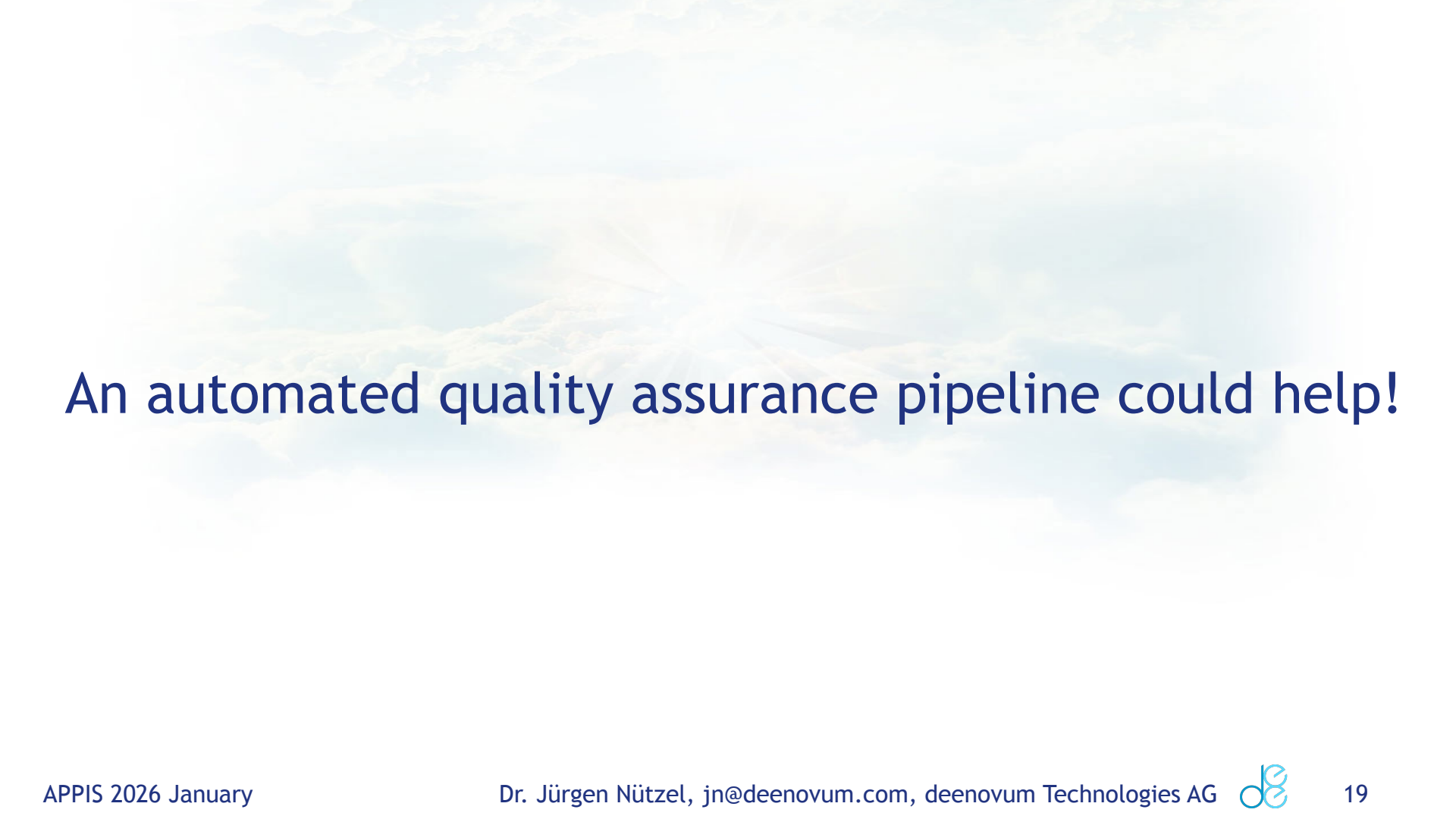
Quality Check

Therefore, we need to conduct a quality check

3

Quantity

Industrial wholesalers may have more than one million products in their assortment



An automated quality assurance pipeline could help!

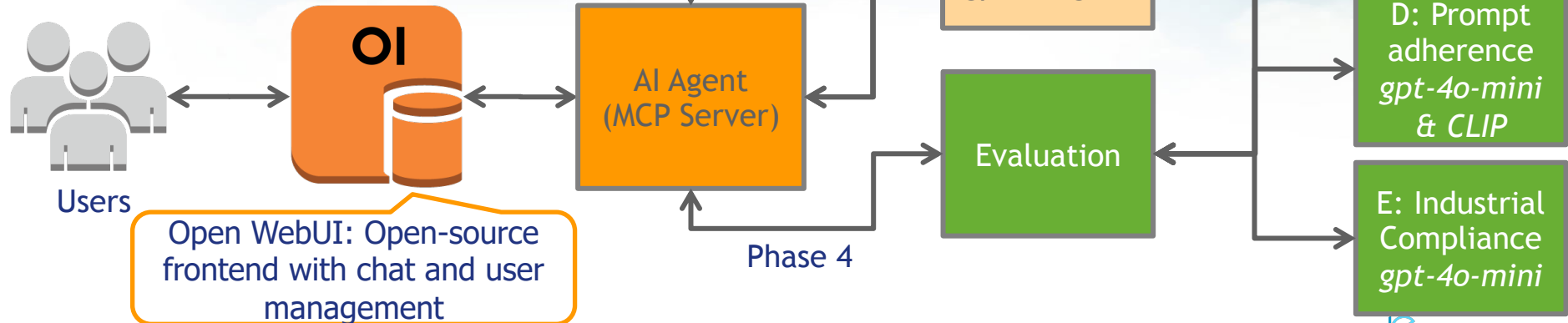
Overall System Architecture

Phase 1: Reads original image and existing product attributes

Phase 2: Enhances prompt by attributes to avoid hallucinating incorrect details

Phase 3: Image generation

Phase 4: Audits created image using five (A-E) distinct metrics



The Evaluation (QA) Metrics

Metric A: Identity (Semantic Consistency)

It calculates the CLIP embedding vector for both images.

It measures the Cosine Similarity between them. If the score is high (near 1.0), the AI preserved the semantic meaning of the object.

Metric B: Structure (SSIM)

It converts images to grayscale and calculates **Structure Similarity Index (SSIM)**.

This measures if textures and lighting gradients were preserved.

Metric C: Edge Integrity (Canny SSIM)

It applies a **Canny Edge Detector** to strip away color and texture, leaving only wireframe outlines.

It compares the wireframes. This ensures the shape of the object hasn't warped.

Metric D: Prompt Adherence

It asks gpt-4o-mini to write a description of the "Before" state and the "After" state.

It uses CLIP to measure the probability that the Edited Image matches the "After" description better than the "Before" description.

Metric E: Industrial Compliance

- Although the previous four metrics would be sufficient in any industry where critical attributes are not important, images of industrial objects must represent their attributes as accurately as possible.
- It sends the Edited Image and the *Ground Truth* attributes to gpt-4o-mini.
- The LLM acts as an industrial inspector, checking for violations (e.g. “Spec says aluminium, image shows rusted iron”)
- It returns a Pass or Fail Boolean.

Results

For the initial tests, we used a threaded rod.



Threaded rod's (Gewindestange) result

Product: Gewindestange M8.
Attributes: Length 1000 mm,
material steel, surface
electrogalvanized.

Original Image



Prompt 1:

“Place this product on a clean factory floor with yellow safety lines in the background.”



- Metric A: 0,8057
- Metric B: 0,6284
- Metric C: 0,7958
- Metric D: 0,9771
- Metric E: **TRUE**. The image shows a threaded rod which matches the specified product type. The appearance is metallic and galvanized, consistent with the material specification. Length cannot be assessed, but shape and material comply.

Prompt 4:

“Change the material of this object to polished brass.”



- Metric A: 0,8367
- Metric B: 0,8198
- Metric C: 0,8346
- Metric D: 0,9846
- Metric E: **FALSE**. Material appearance is metallic gold, which does not match the specified galvanized steel

Prompt 7:

“Add a semi-transparent watermark text 'CONFIDENTIAL' diagonally across the image.”



- Metric A: 0,7167
- Metric B: 0,5034
- Metric C: 0,7305
- Metric D: 0,8643
- Metric E: **TRUE**. The image shows a threaded rod with indications of being M8, steel material, and a galvanized appearance. Length specifications cannot be assessed, but no gross shape mismatch identified.

Bonus Prompt:

“Put the object in the hands of an Instagram Model.”



- Metric A: 0,7261
- Metric B: 0,5241
- Metric C: 0,6978
- Metric D: 0,9361
- Metric E: **TRUE**. The image shows a threaded rod that appears to be metallic and has a galvanized finish, matching the material specification. There is no gross shape mismatch, and length cannot be assessed.

Summary and lessons learned

- LLM can generate industrial product images
- LLMs are not reliable
- We have a quality and quantity problem
- LLMs can also help us with this.
- We have developed a quality assurance pipeline
- Initial experiments give us confidence
- We need to keep an eye on costs and speed

Thank you for your attention

PD Dr-Ing. habil. Jürgen Nützel, CEO, jn@deenovum.com
Fahad Mansuri, fm@deenovum.com

APPIS 2026
8th International
Conference
Las Palmas de Gran
Canaria, Spain
January 19-23, 2026



Find me on LinkedIn