

Generative-AI-assisted knowledge-based augmented exploration in nanophotonics

Masanobu Iwanaga ^a and Keisuke Watanabe ^b

^aResearch Center for Electronic and Optical Materials, National Institute for Materials Science (NIMS), Tsukuba, Japan; ^bResearch Center for Materials Nanoarchitectonics (MANA), National Institute for Materials Science (NIMS), Tsukuba, Japan

ABSTRACT

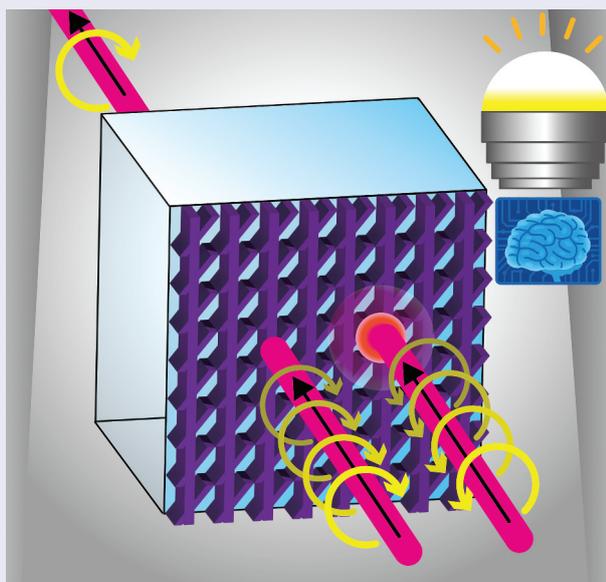
It is generally difficult to obtain valuable scientific findings without sufficient background knowledge or experience. Exploiting generative AI (genAI) and reliable computational method(s) strategically is expected to reduce the difficulty and attain scientific findings beyond human knowledge and experience. Here, we introduce a new approach that initially relies on a substantial amount of scientific literature to eliminate hallucinations, prepares metadata from the literature using genAIs, conducts retrieval augmented generation using a genAI, implements genAI-driven generations of concrete candidates of nanostructures, and evaluates the candidates quantitatively, which contain extended sets of structural parameters, by implementing a scientifically established simulation method. This genAI-assisted nonempirical approach has been applied in the field of nanophotonics, enabling the exploration of single-layer circular dichroic (CD) all-dielectric metasurfaces in telecom bands. We successfully revealed single-layer perfect CD metasurfaces, which have never been attained so far. The perfect CD metasurfaces are thin with subwavelength thickness in the near-infrared bands, are feasible in a semiconductor nanofabrication process, and significantly enlarge degrees of freedom in light-wave manipulations for telecom applications.

ARTICLE HISTORY

Received 18 July 2025
Revised 7 November 2025
Accepted 25 November 2025

KEYWORDS

Generative AI; retrieval augmented generation; nonempirical exploration; nanophotonics; chiral metasurface; circular dichroism



IMPACT STATEMENT

Generative AI assisted nonempirical search for finding new metasurface structures, some of which, consisting only of a single layer, exhibit ideally perfect circular dichroism.

1. Introduction

Artificial intelligence (AI) has drastically expanded its field of application. It began to prove its outstanding capability almost a decade ago by mastering self-

complete board games such as Go, shogi, and chess, which were considered to take tens of years for human brains to be surmounted by computers [1–3]. This breakthrough indicated explicitly that AI reinforced by deep machine learning (ML) would soon excel in

CONTACT Masanobu Iwanaga  iwanaga.masanobu@nims.go.jp  Research Center for Electronic and Optical Materials, National Institute for Materials Science (NIMS), 1-1 Namiki, Tsukuba 305-0044, Japan

© 2025 The Author(s). Published by National Institute for Materials Science in partnership with Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

a large part of human brain activities. Even those who did not pay attention to the board games have recently noticed the power of generative AI (genAI) to master natural (or human) languages [4,5] and image generation [6] through commercialization.

After the breakthrough of mastering the games, a scientific application of genAI has been used to predict protein structures [7–11]. In the predictions, a large number of molecular dynamics simulations for proteins were conducted, forming large metadata that were analyzed/processed using genAIs. One of the keys was that the metadata were not collected from existing literature/data but were prepared by employing a reliable scientific simulation method, molecular dynamics simulation. This substantially expedited the establishment of a scheme for protein structure prediction over several years [12].

To extend the use of genAI to other scientific disciplines, ML-based approaches have progressed over the past decade [13–22]. In fields with existing big data, ML provides candidates for new materials and processes to synthesize novel materials, including molecules. When we turn our eye to other fields, it often occurs that there is insufficient accumulated big data regarding the research fields. In such cases, another approach is necessary to employ genAI; that is, without assuming existing big data, knowledge-based exploration of scientific findings is highly desired in many research fields. We adopted this strategy and applied it in nanophotonics. Figure 1 illustrates the entire process of genAI-based exploration in this study, which starts with input (or published scientific papers), goes through metadata generation, design generation, and electromagnetic simulation, and finally ends with outputting optimized metasurfaces for a specific purpose. Metasurfaces are artificially designed nanostructured surfaces with definite

optical functions. Obviously, the present scheme is different from ML-based optimization approaches that use big data in particular research disciplines, as detailed in Section 2.1.

Nanophotonics requires well-designed artificial nanostructures to realize optical/photonic functions for concrete purposes. Most of the designs have been based on known nanostructures and empirical facts [23–25] or intuitions [26,27]. Metasurfaces are an emerging area in optics and photonics, exhibiting outstanding features and diverse applications, ranging from flatlenses [28–30] to biosensors [31–34]. In nanophotonics, design is one of the key issues in obtaining highly functional photonic units or devices, including metasurfaces; therefore, a large number of trials have been reported thus far [35–40]. These trials for better designs were initially based on a typical structure and conducted local structure modifications to optimize a particular index parameter, such as the transmittance in the waveguides. Thus, they did not create entirely new structures, but found optimized structures for specific purposes. Non-empirical structural searches have been conducted in a few cases [41,42]. In principle, non-empirical searches are possible; however, they require huge computational resources that are difficult to apply to various cases. Therefore, it is highly desirable to establish a new approach to enable us to find suitable designs for diverse applications in nanophotonics.

The structures for large circular dichroism (CD) were often set to be helical [43] in microstructures and larger, probably because ‘chiral’ sounds like ‘helical’ or stimulates a notion connected to something helical in mind. However, such helical structures are impractical for high-precision nanofabrication. Chiral metasurfaces are an intriguing topic in nanophotonics [44–47] and are being explored extensively based on human

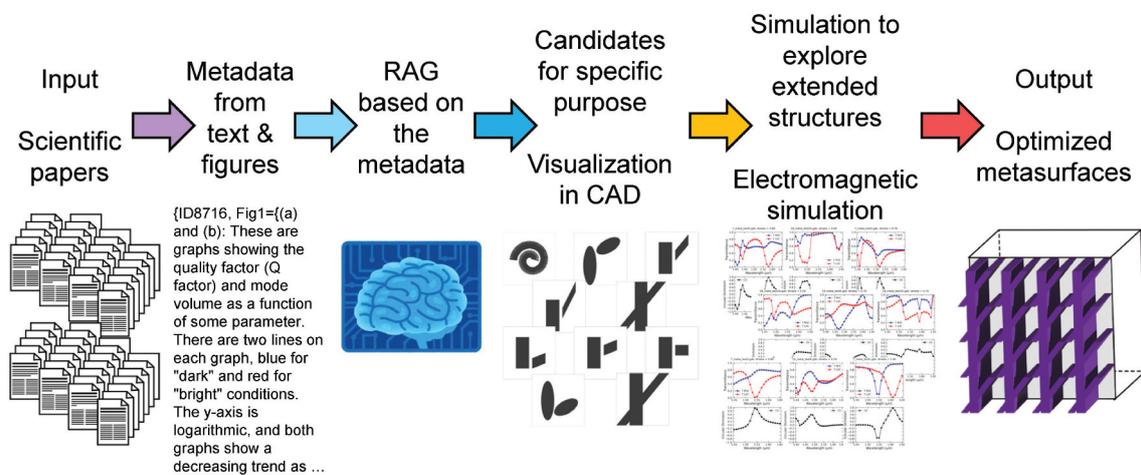


Figure 1. The whole process of genAI-assisted knowledge-based exploration in this study. It begins with published scientific papers, goes through metadata preparation and candidate generation, evaluates extended candidates by electromagnetic simulation, and leads to the identification of optimized nanostructures, which are single-layer metasurfaces in this exploration. RAG and CAD denote the retrieval-augmented generation and computer-aided design, respectively.

knowledge and/or experience. The reported CD responses are associated with small changes in linear optical signals, such as reflection or transmission, of less than 50% [44–47], and are far from perfect CD, satisfying the 0–100% ideal contrast in the optical signals.

We here aim at achieving chiral metasurfaces with perfect CD using the scheme illustrated in Figure 1. If a single-layer nanostructure that exhibits perfect CD is found, it leads to substantial advances in chiral metasurfaces and opens a route to a wide range of applications in light-wave manipulation. In this study, we think highly of potential applications, set the wavelength range to be near-infrared (IR) telecom bands, explore CD all-dielectric metasurfaces consisting only of single layers, and exhibit perfect CD metasurfaces that have never been found in numerous trials to date.

2. Methods

2.1. Whole process

The entire process for exploring nanophotonics is illustrated in Figure 1. We implemented this scheme using a cloud service, which was Amazon Web Service (AWS) [48]. In the initial step, relevant published papers in portable document format (PDF) were accumulated in AWS S3 storage. The total number of accumulated papers was greater than 5,000. The retrieval procedures for the metadata of the text and figures are described in Section 2.2.

Based on the metadata, we extracted potential nanostructures in a visualized manner, as described

in Section 2.3; that is, output computer-aided design (CAD) formation on metasurfaces that meet our desired purpose, which was set as CD. This procedure enabled us to focus on particular structures among the numerous metasurface structures that could show CD, making design procedures practical (or design time substantially shorter).

The output CAD information provided us with the potential CD structures. It is necessary to determine good structures in a scientific manner. We implemented a series of large-scale electromagnetic simulations on the cloud service by exploiting high-performance computing (HPC) systems, which are available at AWS EC2. The results were also output as metadata for the cloud storage. This step is described in Section 2.4. Conducting the analysis of the metadata, we obtained the prominent CD metasurfaces described in Section 3.

2.2. Retrieval augmented generation (RAG) on cloud

Figure 2 shows a schematic of retrieval augmented generation (RAG) scheme used in this study. In this study, we intended to create a genAI scheme specific to a particular scientific discipline, which was nanophotonics; therefore, we initially provided a substantial number of published papers of more than 5,000 in PDF as a base set of data regarding the discipline. The PDF papers were collected by ourselves under a criterion that they address nanophotonics containing metamaterials, metasurfaces,

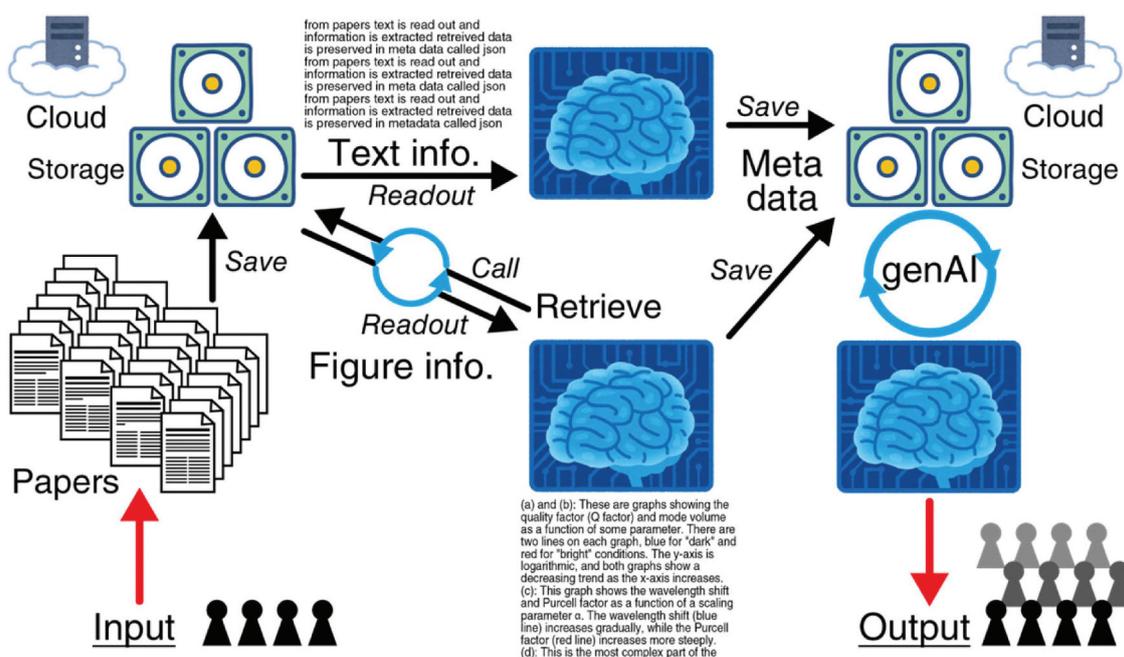


Figure 2. Schematic of RAG on cloud in this study. Based on the relevant data, which were a substantial number of published papers in a specific discipline of nanophotonics, extractions were conducted for the text and files in the given data using two genAIs independently, preparing metadata. A genAI relied only on the metadata and output answers to questions via text input and output.

photonic crystals, nanoparticle photonics, and, in part, optical sciences. There was no other criterion to avoid forming a biased dataset. The base dataset was stored on the cloud storage, Amazon S3 [48].

Two-way extractions from the base data to produce metadata were conducted: one was conducted by extracting text information using a genAI, which was Amazon Titan Text G1 running on AWS [48] and the other was done by extracting figure information using another genAI, which was Anthropic Claude 3.5 Sonnet v2 on AWS. On the one hand, the extraction from text in the PDF files was implemented in a built-in scheme in AWS; that is, the text in the PDF file stored in an AWS S3 directory was read out by the Titan Text G1 and the metadata in json format was stored in a designated directory in the Amazon S3. This scheme was provided as Knowledge Base in Amazon Bedrock [48]. On the other hand, scientific papers usually have several figures that convey important information about the content. Therefore, it is necessary to include figure information in metadata. However, information extraction from figures in PDF files was not default at the moment when we developed this exploration scheme; only a figure file was retrieved in the Claude 3.5 Sonnet.

A large number of figures in thousands of PDF files was needed to be automatically read. For this purpose, we built a pipeline to extract abundant figure information on the cloud (AWS) by integrating various services. The procedure is as follows:

- (1) A series of codes was prepared and saved in CodeCatalyst.
- (2) The codes were stacked in CloudFormation and deployed in Elastic Container Registry (ECR).
- (3) The ECR sent commands to Lambda, which executes the commands in the codes.
- (4) The Lambda picked up a PDF file from the directory of AWS S3 and sent it to a genAI, the Claude 3.5 Sonnet, to read out the figure information.
- (5) The genAI read figures, including the captions and related text, and output the metadata for each figure. The metadata in the json format was output by the Claude 3.5 Sonnet was stored in a designated directory in Amazon S3.
- (6) Returned to the step (4) above and repeated the execution until all the PDF files have been read.

Thus, the data input, the extractions from the text and figures, and saving the meta- data were implemented on a cloud system in a closed manner, as illustrated in Figure 2. One of the readout results by the genAI is shown in Appendix A, indicating that the generated descriptions are more informative than the original caption and that the precision of the readout is fairly high. We note that it took approximately 20 s for the

genAI to extract a set of information on a typical figure; accordingly, approximately 3 min were typically spent on a paper with 8–10 figures.

We set the metadata as *knowledge* and used the Claude 3.5 Sonnet as a genAI that outputs answers for the queries raised. This is the RAG scheme used in this study. We examined the precision of the RAG scheme by evaluating the answers to the knowledge-based genAI. Four nanophotonics experts contributed to the queries and asked 200 questions. From an expert viewpoint, approximately 90% of the answers were correct and/or reasonable. The remaining 10% of the answers were classified as failures; some of the failures were due to too specialized questions that may have been solved by only one person in the world. We provide some of question-and-answer (Q&A) examples in Appendix B.

One of the widely known issues in genAI is hallucinations. We sometimes encountered obvious hallucinations of commercial genAIs regarding questions on nanophotonics. However, we did not find any hallucinations in the present knowledge-based genAI, probably because the RAG was strictly based on the metadata we prepared. When the questions addressed out-of-scope issues, the answers were faithfully that the RAG has no information on the questions; in contrast, commercial genAI tends to answer questions even if they do not have the answers in reality, thereby giving rise to hallucinations with a high probability. Thus, the present scheme in Figure 2 did function as expected.

2.3. Flow for visualization

The metadata obtained in Section 2.2 enabled us to conduct questions and answers in text, as usual in genAIs. To extract more useful information from the knowledge-based genAI, we prepared a scheme to output visualized nanostructures as CAD files.

The Bedrock in AWS has a function called Flow, which enables the generation of end-to-end genAI workflows by connecting graphical nodes in a visual interface. We built a multi-step workflow that includes

- (1) query for desired metasurface structures using RAG,
- (2) generation of a Python script to create a CAD file that visualizes the text-based output obtained from RAG, and
- (3) upload of the generated script file to cloud storage (Amazon S3). Figure 3 shows the detailed flow of this process.

The input to the flow was provided in the form of a json object containing both the RAG query to find the desired structure and the output destination path in Amazon S3 as follows:

```
{'input': "Can you suggest an asymmetric
chiral metasurfaces or
```

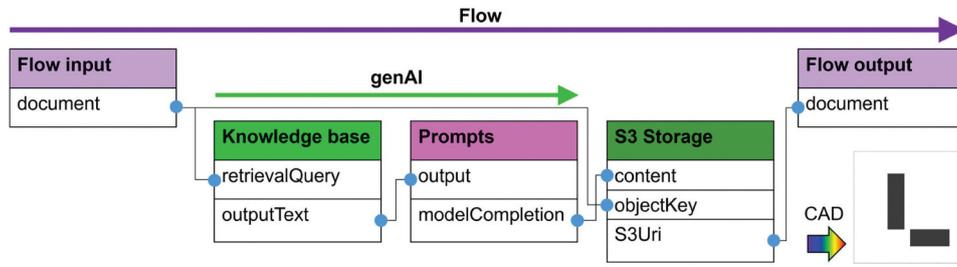


Figure 3. Flow diagram to produce CAD files of metasurface unit cells by employing knowledge-based genAI. The CAD files were output in the GDS format via the Gdstk library [49] in Python. For natural language input/output, another genAI was used to execute the flow.

metamaterials that have a large circular dichroism?
Please provide the structural parameters and dimensions as much detail as possible". ,

‘s3key’: ‘outputs/test.py’}.

We repeated this procedure and stored tens of CAD files for further steps, as described in Section 2.4.

In Figure 3, the sequence of steps in the Flow is shown. The input string was first passed to the Knowledge Base, where Titan Text Embeddings v2 was used as the embedding AI model. The Knowledge Base implements the RAG provided in the AWS Bedrock. The resulting retrieved data included detailed structural information, such as the geometry and dimensions of the chiral metastructures expected to exhibit large CD. Next, this output was sent to Prompts, run by the Claude 3.5 Sonnet, to generate a Python script that creates a CAD file in GDSII format. This Python script was generated using the Gdstk (GDSII Tool Kit) Python library [49], and subsequently saved to the specified S3 bucket location defined by s3key. Here, prompt engineering was

employed to guide the generative model to create only 2D unit cell structures exhibiting optical resonances in a near-IR range from 1.4 to 1.7 μm. The GDS file was successfully created by executing Python script. An example of the resulting GDS file is shown in Figure 3 (bottom right). We confirmed that the structural information retrieved using the RAG-based genAI tool was effectively visualized and exported in a standard CAD format.

2.4. Nonempirical exploration for nanophotonic CD structures

Figure 4(a) shows the structural and optical configurations of interest. The incident plane wave propagates from the substrate side, transmits to the metasurface layer, and reaches the air side, where the transmitted light is observed. Right- and left-handed circular (RHC and LHC, respectively) polarizations are defined for the incident wavevector k_{in} of normal incidence to the metasurface, as shown in Figure 4(a). The electromagnetic simulation was conducted using a finite-difference

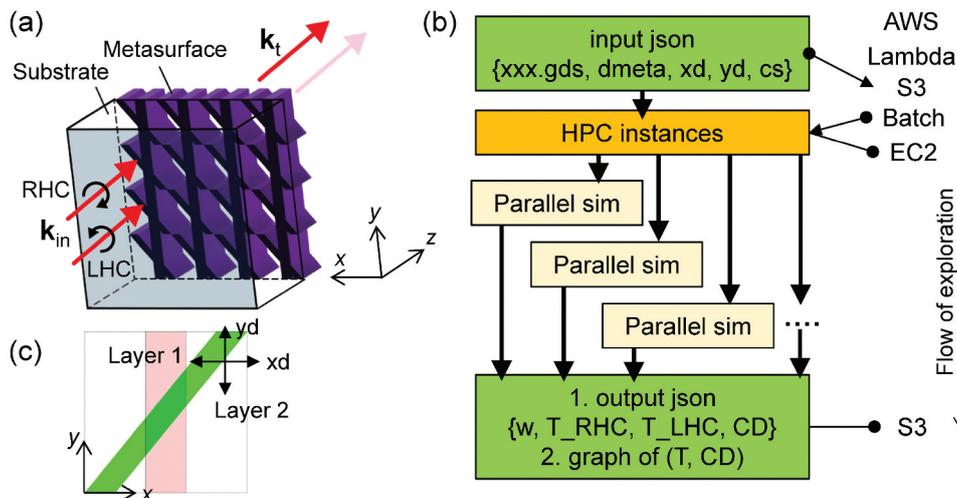


Figure 4. Structural search for single-layer CD metasurfaces. (a) Optical configuration for electromagnetic simulations. Incidence with wavevector k_{in} propagates from the substrate to the air with transmitting the metasurface. (b) Diagram of the structural search implementation. HPC and sim denote high-performance computing and simulation, respectively. (c) xy -section-view unit cell in used in the structural search, composed of layers 1 (pale red) and 2 (green). The position of the object(s) in the layer 2 was varied in accordance with the offset values x_d and y_d that were given in the input json file in (b).

time-domain (FDTD) method package, named MEEP [50,51], which can run on HPC Amazon EC2 instances operated by a Linux system, Ubuntu 21.04 or later.

We confirmed that the FDTD package outputs simulated results in good agreement with those obtained by rigorous coupled-wave analysis (RCWA) combined with the scattering matrix algorithm [52,53]. It has been well established that the RCWA is a reliable method to show good agreements with experimental data regarding linear optical responses, such as transmission and reflection, of dielectric nanostructures including metasurfaces and photonic crystals [54–57].

The scheme implemented in the simulation-based exploration is shown in Figure 4(b). Input information on the structural exploration was provided in a text-based json file, which was associated with a base CAD file in GDSII format. Nanostructures were assumed to be made of Si, which was set to have a representative refractive index of 3.46 in the wavelength range of the present interest, and were defined in the unit cell, being placed in CAD layer 1 (pale red) or 2 (green), as shown in Figure 4(c). The element(s) in the layer 2 is moved with following the values of ‘xd’ and ‘yd’ in the json file; furthermore, the thickness of the metasurface was varied by ‘dmeta’. In addition, the dimension of the unit cell was changed by a factor for crystalline scale ‘cs’; the default was 1.0, and the modified values (e.g. 0.9 and 0.8) can be set in accordance with the change (e.g. respective 10% and 20% reduction) in the size of the unit cell. These structural parameters correspond to the structural modifications of the metasurface. Thus, metasurfaces were explored to find definite CD features in a 3D manner.

For one base CAD file, hundreds or more implementations were set in the json file; therefore, it was crucial to conduct multi-parallel implementations (MPIs). To realize MPIs, we built a pipeline to start and set up tens of Amazon EC2 instances simultaneously for the electromagnetic simulations. As shown on the right-hand side of Figure 4(b), the initial command was executed to read the json file using AWS Lambda, the command called Batch that started tens of EC2 instances in an optimized manner and set up MEEP on the instances, and the MPIs were executed.

After running the MPI structural exploration, the output json files were saved in a directory of Amazon S3. We briefly analyzed the json files by using Glue and Athena in AWS, which are platforms to access and analyze data on the S3, and found the metasurfaces exhibiting the best CD responses. The electromagnetic simulations also output the transmission and CD spectra in another directory of the S3. Thus, we readily glanced at high-performance CD metasurfaces among the hundreds of simulated metasurfaces. Some of the featured spectra are presented in Section 3.

3. Results and discussion

To evaluate defines degree of CD quantitatively, it is defined such that

$$CD = \frac{T_{RHC} - T_{LHC}}{T_{RHC} + T_{LHC}} \quad (1)$$

where T_{RHC} and T_{LHC} denote the transmittance under RHC and LHC polarizations, respectively. In the simulations, the transmittance was normalized, taking the values of $[0, 1]$. From the definition in Equation (1), the degree of CD takes the values of $[-1, 1]$. Ideal degree of CD is realized under the condition of $\{T_{RHC} = 1 \text{ and } T_{LHC} = 0\}$ or $\{T_{RHC} = 0 \text{ and } T_{LHC} = 1\}$, resulting in $CD = 1$ or -1 , respectively.

A set of representative results obtained through the exploration is shown in Figure 5. We show three series of CD metasurfaces, each of which has a unit cell structure, transmittance spectra, and degree of CD.

In Figure 5(a), the unit cell includes two crossing bars of Si (gray), set to a dimension of $0.8 \times 0.8 \mu\text{m}^2$ in the xy plane. The height of the unit cell was set to $0.7 \mu\text{m}$ along the z axis. Note that, because the unit cell arrays periodically, the vertical bar is connected to the vertical bar at the top and bottom edges of the unit cell, while the oblique bar is not connected with a finite contact length to any bar in the nearest next-neighbor unit cells. In Figure 5(b), the transmittance spectra under RHC and LHC polarizations, which were computed using MEEP, are shown with blue and red lines with closed circles, respectively. Evidently, a distinct contrast of transmittance appears at $1.510\text{--}1.525 \mu\text{m}$, which realizes almost perfect CD (i.e. $CD \approx 1$, $T_{RHC} \approx 1$, and $T_{LHC} \approx 0$), as shown in Figure 5(c). These results exhibit a perfect CD band realized by the *cross-bar* metasurface, as shown in Figure 5(a).

The perfect CD response in Figure 5(b,c) is a direct consequence of the distinct optical responses dependent on circular polarization. The metasurface in Figure 5(a) exhibits circular-polarization-dependent high-contrast transmission at $1.510\text{--}1.525 \mu\text{m}$, that is, almost 100% and 0% transmittance under RHC and LHC polarizations, respectively. These results show that a high transmission mode is induced only under RHC polarization and a perfect light absorption mode occurs only under LHC polarization, resulting in zero transmission. Both modes are not usually induced, which is confirmed by the result that the transmittance in most of the wavelength range substantially deviates from 100% and 0%, which is a typical optical response. We examined the underlying mechanisms of the two modes from resonant electromagnetic fields and found a light-guiding mode appears in the cross-bar structure under RHC polarization and that a light-localizing mode is induced around the cross-bar structure under LHC

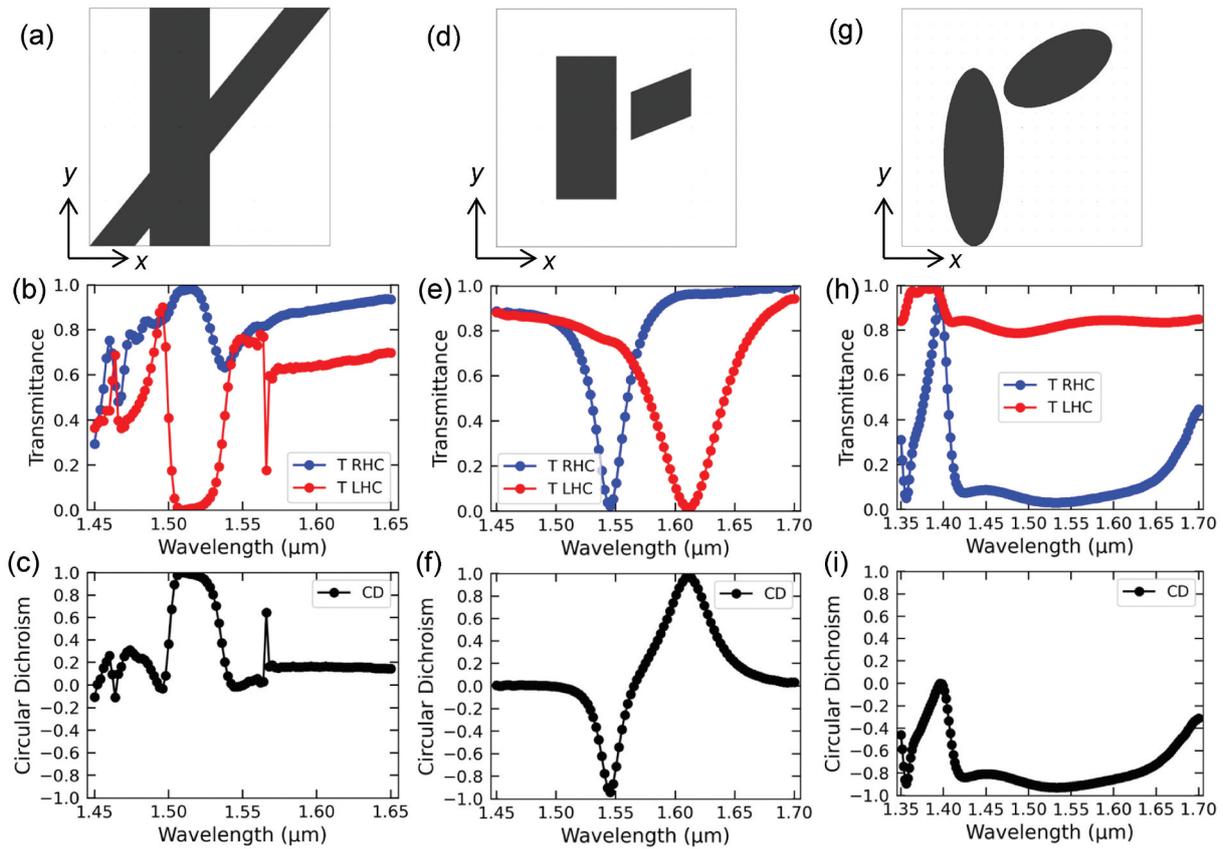


Figure 5. Representative results attained through the augmented exploration. (a) Unit cell of a CD metasurface, where cross-bar Si nanostructure (gray) is shown with gray in an xy -section view. (b) Normalized transmittance spectra of the metasurface in (a) under RHC (blue line with closed circles) and LHC (red line with closed circles) polarizations, respectively. (c) CD spectrum (Black line with closed circles) shows an almost perfect CD response at 1.510–1.525 μm , evaluated by transmittance in (b) and Equation (1). (d) Unit cell of a CD metasurface comprising two-block Si nanostructure (gray), shown in a similar manner to (a). (e) Normalized transmittance spectra of the metasurface in (d) under RHC (blue) and LHC (red) polarizations, presented in a similar manner to (b). (f) CD spectrum (Black) exhibits a feature that the degree of CD takes +1 and -1 , evaluated by transmittance in (e) and Equation (1). (g) Unit cell of a CD metasurface composed of two elliptic Si nanostructures (gray), shown in a similar manner to (a). (h) Normalized transmittance spectra of the metasurface in (g) under RHC (blue) and LHC (red) polarizations, presented in a similar manner to (b). (i) CD spectrum (Black) evaluated by transmittance in (h) and Equation (1) shows a broadband large degree of CD, including the telecom C band. The specific dimensions of the unit cells in (a), (d), and (g) are described in the text.

polarization, resulting in the perfect light absorption. Further detailed analyses of the optical resonances will be reported elsewhere.

Figure 5(d) shows a unit cell of two-block Si nanostructures (gray). The dimension of the unit cell was $0.8 \times 0.8 \mu\text{m}^2$ in the xy plane and the height was $0.9 \mu\text{m}$ along the z axis. This aspect ratio is feasible in deep Si reactive-ion etching [58,59]. Transmittance spectra show two prominent dips at 1.545 and 1.620 μm under RHC and LHC polarizations, respectively, in Figure 5(e), which result in almost perfect degree of CD in Figure 5(f). This CD feature is unique because CD varies over the full range from -1 to $+1$. Thus, the *two-block* metasurface functioned as a unique CD device in the near-IR range.

In Figure 5(g), a unit cell of a pair of two elliptic Si nanostructures with low symmetry is shown, being set to a dimension of $0.8 \times 0.8 \mu\text{m}^2$ in the xy plane and of $0.9 \mu\text{m}$ in height along the z axis. A difference from the structures in Figure 5(a,d) is that the elliptic pair has

a nanogap between them. Transmittance spectra exhibit a broadband response of high T_{LHC} and low T_{RHC} at 1.420–1.620 μm , resulting in a broadband large CD that satisfies with $|CD| \geq 0.8$; in particular, the CD satisfies with $|CD| > 0.9$ at 1.505–1.565 μm . To the best of our knowledge, the broadband CD features realized by the very thin single-layer structure have not been found in the visible and near-IR ranges. The feature indicates that there is not a definite resonant wavelength for any circular polarization, being considered to come from non-resonant dichroic states and functioning as a transmission-selective metamaterial for circular polarizations. The details of physical optics will be examined elsewhere.

Further outputs of the structural exploration are shown in Figure 6. In Figure 6(a), the role of the smaller Si nanoblock was tested, which has a rectangular shape, in contrast to the inclined shape in Figure 5(d). The area dimensions in the smaller nanoblocks in Figures 5(d) and 6(a) were the same. The larger nanoblocks were identical.

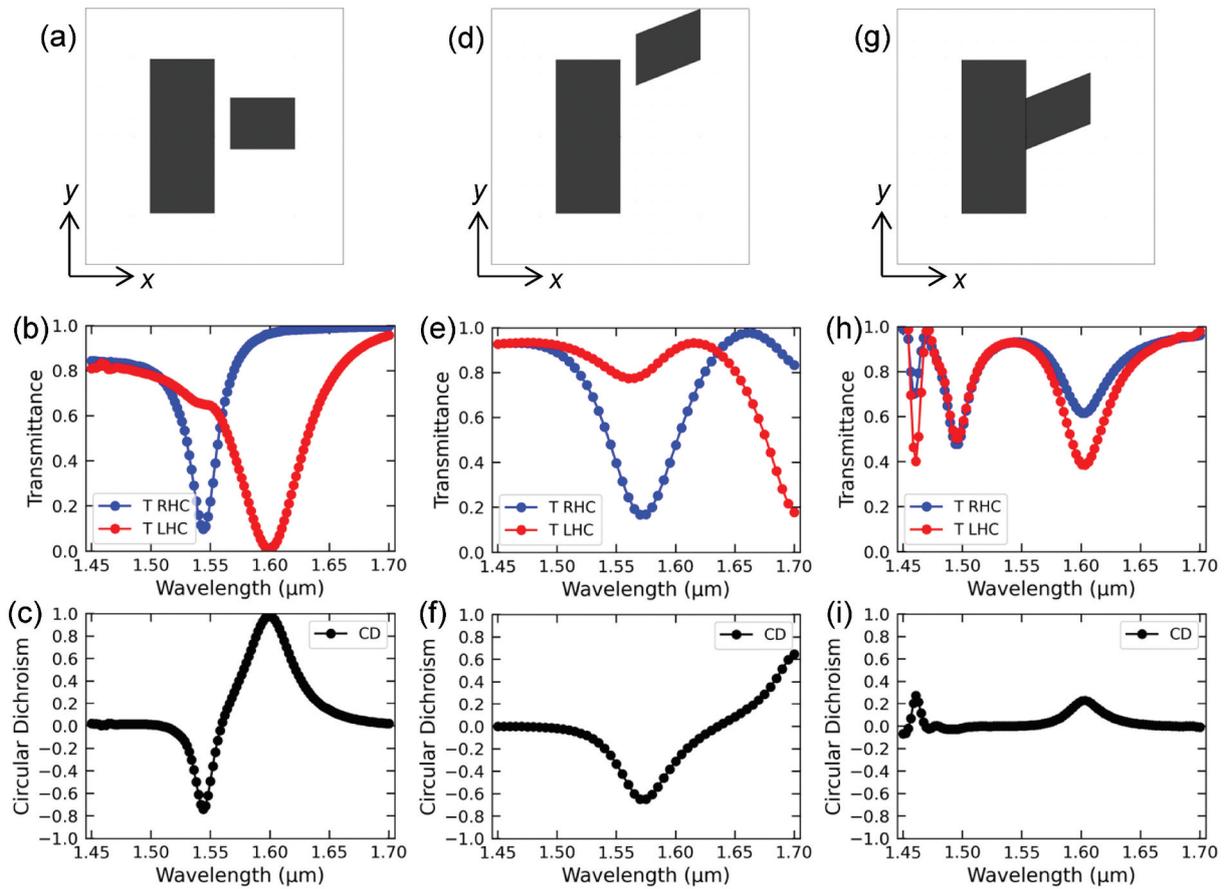


Figure 6. Further results in the exploration of CD metasurfaces. (a) Unit cell of a perfect CD metasurface comprising two rectangular Si blocks (gray) with a 50 nm gap. (b) Normalized transmittance spectra under RHC and LHC polarizations, shown with blue and red lines with circles, respectively. (c) CD spectrum (Black) shows a perfect CD response at 1.60 μm. (d) Unit cell of a CD metasurface, which is a modified structure of Figure 5(d), by changing the value of the y -offset y_d (gray). (e) Transmittance spectra of the metasurface in (d), shown similarly to (b). The transmittance contrast becomes lower than the transmittance shown in Figure 5(e). (f) CD spectrum (Black) shows a reduced degree of CD compared with that in Figure 5(f). (g) Unit cell of a CD metasurface, which is a modified structure of Figure 5(d), by changing the values of x_d and y_d . All Si nanoblocks were connected to form a connected structure (gray). (h) Transmittance spectra of the metasurface in (g), as shown similarly to (b). (i) CD spectrum (Black) shows a further reduced degree of CD compared with (f).

In addition, the xyz dimension of the unit cell was set to the same as that shown in Figure 5(d). Transmittance spectra in Figure 6(b) are qualitatively similar to those in Figure 5(e), while they are quantitatively different; in the former, the transmittance under RHC polarization (blue) does not go down to zero and therefore the degree of CD was not perfect, reducing such that $|CD| < 0.8$ in Figure 6(c). Thus, the design in Figure 5(d) was found to be better than Figure 6(a) in terms of the degree of CD.

In Figure 6(d), a larger y -offset, y_d , than that in Figure 5(d) was tested. The dimensions of the unit cell in the xy plane and the height of the Si nanoblocks are identical to those shown in Figure 5(d). It turned out that the contrast of transmittance became lower than that in Figure 5(e); consequently, the resultant degree of CD was reduced in the regime of $|CD| \leq 0.8$, as seen in Figure 6(f).

Structural modifications often reduce the degree of CD. Figure 6(g) shows a connected Si nanoblock (gray) in the unit cell, which is a modification of the unit cell shown in Figure 5(d). Transmittance spectra

in Figure 6(h) present substantial overlap under RHC and LHC polarizations. Consequently, the degree of CD was reduced significantly, compared to that shown in Figure 5(f).

Overall, the perfect or broadband CD responses in Figures 5 and 6 were attained with good combinations of the base structure defined by the CAD file and the structural parameters $\{d_{\text{meta}}, x_d, y_d, c_s\}$ in Figure 4. The single-layer perfect CD metasurfaces were very difficult to reach without substantial suggestions from the present genAI-assisted exploration. Thus, we believe that there is no surprise regarding the fact that, despite the extensive effort devoted to studies on CD metasurfaces, no one has succeeded in finding such single-layer perfect CD metasurfaces.

4. Conclusions

We have introduced a new scheme for nonempirical exploration for nanophotonic metasurfaces, which was assisted by genAIs and based on the

relevant scientific papers. The genAIs were used in several roles, such as text input/output interface, extraction from text and figures, augmented generation of graphical output in the CAD format, and text output in json format. Starting from the genAI outputs on the CD metasurfaces, we implemented augmented electromagnetic simulations for the CD metasurfaces by providing extended sets of structural parameters. Consequently, we have achieved single-layer perfect or broadband CD all-dielectric metasurfaces working in near-IR telecom bands, which have not been attained from conventional design approaches based on human knowledge and experience. The uncovered perfect or broadband CD metasurfaces with less than 1 μm thickness can contribute to highly efficient light-wave manipulations at the optical telecom bands, functioning as novel elements for ultrathin optical devices.

Acknowledgements

We acknowledge Ya-Lun Ho and Tetsuyuki Ochiai (NIMS) for their contribution to an early stage evaluation of the RAG based on published papers on nanophotonics, and Yusuke Ogiwara (TPC, LLC) for assistance in building the pipelines in AWS.

Disclosure statement

We note that a part of this study was filed as a Japanese Patent, 2025-142888.

Funding

This study was partially supported by the NIMS AI-use Research Project [Generative AI-use Research Project]. MI acknowledges JSPS KAKENHI Grant Number [JP24K01389] for its financial support.

Notes on contributor

Masanobu Iwanaga initiated and managed this study, which included the preparation of the knowledge-based RAG on the cloud system, electromagnetic simulations, and execution of the augmented electromagnetic simulations. WK built the flow for visualization and partially conducted the augmented electromagnetic simulations. All authors contributed to the writing of this manuscript and agreed with the final manuscript for submission.

ORCID

Masanobu Iwanaga  <http://orcid.org/0000-0002-8930-6940>

Keisuke Watanabe  <http://orcid.org/0000-0002-4285-2135>

References

- [1] Silver D, Huang A, Maddison CJ, et al. Mastering the game of go with deep neural networks and tree search. *Nature*. 2016;529(7587):484–489. doi: 10.1038/nature16961
- [2] Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of go without human knowledge. *Nature*. 2017;550(7676):354–359. doi: 10.1038/nature24270
- [3] Silver D, Hubert T, Schrittwieser J, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*. 2018;362(6419):1140–1144. doi: 10.1126/science.aar6404
- [4] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *Adv Neural Inf Process Syst*. 2017;30:5998–6008.
- [5] Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst*. 2020;33:1877–1901.
- [6] Rombach R, Blattmann A, Lorenz D, et al. High-resolution image synthesis with latent diffusion models. *Proc IEEE/CVF Conf Comput Vision Pattern Recogn*. 2022;1(1):10684–10695.
- [7] Wang S, Sun S, Li Z, et al. Accurate de novo prediction of protein contact map by ultra-deep learning model. *PLOS Comput Biol*. 2017;13(1):e1005324. doi: 10.1371/journal.pcbi.1005324
- [8] Senior AW, Evans R, Jumper J, et al. Improved protein structure prediction using potentials from deep learning. *Nature*. 2020;577(7792):706–710. doi: 10.1038/s41586-019-1923-7
- [9] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583–589. doi: 10.1038/s41586-021-03819-2
- [10] Tunyasuvunakool K, Adler J, Wu Z, et al. Highly accurate protein structure prediction for the human proteome. *Nature*. 2021;596(7873):590–596. doi: 10.1038/s41586-021-03828-1
- [11] Baek M, DiMaio F, Anishchenko I, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. 2021;373(6557):871–876. doi: 10.1126/science.abj8754
- [12] The Nobel Prize in Chemistry. 2024 [cited 2025 Jul 14]. Available from: <https://www.nobelprize.org/prizes/chemistry/2024/summary/>
- [13] Jain A, Ong SP, Hautier G, et al. Commentary: the materials project: a materials genome approach to accelerating materials innovation. *APL Mater*. 2013;1(1):011002. doi: 10.1063/1.4812323
- [14] Liu Y, Zhao T, Ju W, et al. Materials discovery and design using machine learning. *J Materiomics*. 2017;3(3):159–177. doi: 10.1016/j.jmat.2017.08.002
- [15] Butler KT, Davies DW, Cartwright H, et al. Machine learning for molecular and materials science. *Nature*. 2018;559(7715):547–555. doi: 10.1038/s41586-018-0337-2
- [16] Lookman T, Balachandran PV, Xue D, et al. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput Mater*. 2019;5(1):21. doi: 10.1038/s41524-019-0153-8
- [17] Schmidt J, Marques MRG, Botti S, et al. Recent advances and applications of machine learning in solid-state materials science. *npj Comput Mater*. 2019;5(1):83. doi: 10.1038/s41524-019-0221-0

- [18] Himanen L, Geurts A, Foster AS, et al. Data-driven materials science: status, challenges, and perspectives. *Adv Sci*. 2019;6(21):1900808. doi: 10.1002/adv.201900808
- [19] Schleder GR, Padilha ACM, Acosta CM, et al. From DFT to machine learning: recent approaches to materials science—a review. *J Phys Mater*. 2019;2(3):032001. doi: 10.1088/2515-7639/ab084b
- [20] Wang T, Zhang C, Snoussi H, et al. Machine learning approaches for thermoelectric materials research. *Adv Funct Mater*. 2020;30(5):1906041. doi: 10.1002/adfm.201906041
- [21] Chen C, Zuo Y, Ye W, et al. A critical review of machine learning of energy materials. *Adv Energy Mater*. 2020;10(8):1903242. doi: 10.1002/aenm.201903242
- [22] Choudhary K, DeCost B, Chen C, et al. Recent advances and applications of deep learning methods in materials science. *npj Comput Mater*. 2022;8(1):59. doi: 10.1038/s41524-022-00734-6
- [23] Iwanaga M. Ultracompact waveplates: approach from metamaterials. *Appl Phys Lett*. 2008;92(15):153102. doi: 10.1063/1.2909500
- [24] Iwanaga M. Subwavelength orthogonal polarization rotator. *Opt Lett*. 2010;35(2):109–111. doi: 10.1364/OL.35.000109
- [25] Iwanaga M. Photonic metamaterials: a new class of materials for manipulating light waves. *Sci Technol Adv Mater*. 2012;13(5):053002. doi: 10.1088/1468-6996/13/5/053002
- [26] Iwanaga M. Subwavelength electromagnetic dynamics in stacked complementary plasmonic crystal slabs. *Opt Express*. 2010;18(15):15389–15398. doi: 10.1364/OE.18.015389
- [27] Kurosawa H, Choi B, Sugimoto Y, et al. High-performance metasurface polarizers with extinction ratios exceeding 12000. *Opt Express*. 2017;25(4):4446–4455. doi: 10.1364/OE.25.004446
- [28] Yu N, Capasso F. Flat optics with designer metasurfaces. *Nat Mater*. 2014;13(2):139–150. doi: 10.1038/nmat3839
- [29] Genevet P, Capasso F, Aieta F, et al. Recent advances in planar optics: from plasmonic to dielectric metasurfaces. *Optica*. 2017;4(1):139–152. doi: 10.1364/OPTICA.4.000139
- [30] Liang H, Martins A, Borges BHV, et al. High performance metalenses: numerical aperture, aberrations, chromaticity, and trade-offs. *Optica*. 2019;6(12):1461–1470. doi: 10.1364/OPTICA.6.001461
- [31] Tseng ML, Jahani Y, Leitis A, et al. Dielectric metasurfaces enabling advanced optical biosensors. *ACS Photonics*. 2021;8(1):47–60. doi: 10.1021/acsp Photonics.0c01030
- [32] Iwanaga M. Productive biosensing techniques empowered by all-dielectric metasurfaces. *Front Bioeng Biotechnol*. 2025;12:1484638. doi: 10.3389/fbioe.2024.1484638
- [33] Iwanaga M, Hu Q, Tang Y. Metasurface biosensors: status and prospects. *Appl Phys Rev*. 2025;12(2):021305. doi: 10.1063/5.0253333
- [34] Iwanaga M. Single-microRNA detection on high-selectivity metasurface fluorescence biosensors. *ACS Nano*. 2025;19(44):38841–38848. doi: 10.1021/acsnano.5c15853
- [35] Molesky S, Lin Z, Piggott AY, et al. Inverse design in nanophotonics. *Nat Photonics*. 2018;12(11):659–670. doi: 10.1038/s41566-018-0246-9
- [36] Campbell SD, Sell D, Jenkins RP, et al. Review of numerical optimization techniques for meta-device design. *Opt Mater Express*. 2019;9(4):1842–1863.
- [37] Li Z, Pestourie R, Lin Z, et al. Empowering metasurfaces with inverse design: principles and applications. *ACS Photonics*. 2022;9(7):2178–2192. doi: 10.1021/acsp Photonics.1c01850
- [38] An S, Zheng B, Julian M, et al. Deep neural network enabled active metasurface embedded design. *Nanophotonics*. 2022;11(17):4149–4158. doi: 10.1515/nanoph-2022-0152
- [39] Ueno A, Hu J, An S. AI for optical metasurface. *npj Nanophoton*. 2024;1(1):36. doi: 10.1038/s44310-024-00037-2
- [40] Kim M, Park H, Shin J. Nanophotonic device design based on large language models: multilayer and metasurface examples. *Nanophotonics*. 2025;14(8):1273–1282. doi: 10.1515/nanoph-2024-0674
- [41] Iwanaga M. Optically deep asymmetric one-dimensional metallic grooves: genetic algorithm approach. *J Opt Soc Am B*. 2009;26(5):1111–1118. doi: 10.1364/JOSAB.26.001111
- [42] Iwanaga M. Non-empirical large-scale search for optical metasurfaces. *Nanomaterials*. 2020;10(9):1739. doi: 10.3390/nano10091739
- [43] Gansel JK, Thiel M, Rill MS, et al. Gold helix photonic metamaterial as broadband circular polarizer. *Science*. 2009;325(5947):1513–1515. doi: 10.1126/science.1177031
- [44] Zhang X, Liu Y, Han J, et al. Chiral emission from resonant metasurfaces. *Science*. 2022;377(6611):1215–1218. doi: 10.1126/science.abq7870
- [45] Chen Y, Deng H, Sha X, et al. Observation of intrinsic chiral bound states in the continuum. *Nature*. 2023;613(7944):474–478. doi: 10.1038/s41586-022-05467-6
- [46] Gryb D, Wendisch FJ, Aigner A, et al. Two-dimensional chiral metasurfaces obtained by geometrically simple meta-atom rotations. *Nano Lett*. 2023;23(19):8891–8897. doi: 10.1021/acsnanolett.3c02168
- [47] Choi M, Alù A, Majumdar A. Observation of photonic chiral flatbands. *Phys Rev Lett*. 2025;134(10):103801. doi: 10.1103/PhysRevLett.134.103801
- [48] Amazon Web Service. 2025. [cited 2025 Jul 14]. Available from: <https://aws.amazon.com>
- [49] Gdstk. 2025. [cited 2025 Jul 14]. Available from: <https://github.com/heitzmann/gdstk>
- [50] Oskooi AF, Roundy D, Ibanescu M, et al. Meep: a flexible free-software package for electromagnetic simulations by the FDTD method. *Comput Phys Commun*. 2010;181:687–702.
- [51] MEEP. 2025. [cited 2025 Jul 14]. Available from: <https://meep.readthedocs.io/en/latest>
- [52] Li L. Formulation and comparison of two recursive matrix algorithms for modeling layered diffraction gratings. *J Opt Soc Am A*. 1996;13(5):1024–1035. doi: 10.1364/JOSAA.13.001024
- [53] Li L. New formulation of the Fourier modal method for crossed surface-relief gratings. *J Opt Soc Am A*. 1997;14(10):2758–2767. doi: 10.1364/JOSAA.14.002758
- [54] Iwanaga M. Plasmonic resonators: fundamentals, advances, and applications. Singapore: Pan Stanford Publishing; 2016. doi: 10.1201/9781315364711

- [55] Iwanaga M. All-dielectric metasurfaces with high-fluorescence-enhancing capability. *Appl Sci.* **2018**;8(8):1328. doi: [10.3390/app8081328](https://doi.org/10.3390/app8081328)
- [56] Iwanaga M. All-dielectric metasurface fluorescence biosensors for high-sensitivity antibody/antigen detection. *ACS Nano.* **2020**;14(12):17458–17467. doi: [10.1021/acsnano.0c07722](https://doi.org/10.1021/acsnano.0c07722)
- [57] Hu Q, Iwanaga M, Tang Y. Metasurface platform incorporating aggregation induced emission based biosensor for enhanced human serum albumin detection. *Adv Opt Mater.* **2024**;12(23):2400868. doi: [10.1002/adom.202400868](https://doi.org/10.1002/adom.202400868)
- [58] Sumitomo Precision Products Co, Ltd. **2025**. [cited 2025 Nov 4]. Available from: <https://www.spp.co.jp/en/infinity/reason/si/>
- [59] Samco-UCP Ltd. **2025**. [cited 2025 Nov 4]. Available from: <https://samco-ucp.com/processes/siliconetching/deepsietchingboschprocess/>
- [60] Liang Y, Tsai DP, Kivshar Y. From local to nonlocal high-Q plasmonic metasurfaces. *Phys Rev Lett.* **2024**;133(5):053801. doi: [10.1103/PhysRevLett.133.053801](https://doi.org/10.1103/PhysRevLett.133.053801)
- [61] Tonkaev P, Toftul I, Lu Z, et al. Nonlinear chiral metasurfaces based on structured van der Waals materials. *Nano Lett.* **2024**;24(34):10577–10582. doi: [10.1021/acs.nanolett.4c02765](https://doi.org/10.1021/acs.nanolett.4c02765)
- [62] Watanabe K, Iwanaga M. Nanogap enhancement of the refractometric sensitivity at quasi-bound states in the continuum in all-dielectric metasurfaces. *Nanophotonics.* **2023**;12(1):99–109. doi: [10.1515/nanoph-2022-0565](https://doi.org/10.1515/nanoph-2022-0565)

Appendices

Appendix A. Retrieval of figure contents by genAI

Figure A1 shows an example of a figure extracted by the knowledge-based genAI noted in Section 2.2. This figure addresses light confinement effect in plasmonic structures (or metallic nanostructures); parts (a)–(c) comprise plots of physical quantities, such as quality (Q) factor, mode volume, and wavelength, for scaling parameter α , and part (d) shows nonlocal and local features of the plasmonic resonances, including graphical radiation patterns. Thus, this figure is a complicated mixture of the physical quantity information and several display items. The followings are original figure captions, quoted from the literature [60].

Eigenmode analysis at the Γ point of VSRR metasurfaces. (a)–(c) The Q factor, mode volume, and resonance wavelength dependence on the scaling parameter α for the dark and bright modes. The initial parameters (at $\alpha = 1$) are period $3 \mu\text{m}$, square pillar width $0.4 \mu\text{m}$, and high $1.8 \mu\text{m}$, middle connector height $0.5 \mu\text{m}$, center-to-center distance between pillars $0.8 \mu\text{m}$. (c) Inset: Γ -point position. (d) The radiation patterns and electric fields (E_z) for various α .

We implemented a series of extractions for figures in more than 5,000 published papers. The genAI output a set of descriptions on Figure A1, as follows:

This figure appears to be a comprehensive representation of the properties and behavior of a nanophotonic system, likely related to optical antennas or nanostructures. Let me break down the components: (a) and (b): These are graphs showing the quality factor (Q factor) and mode volume as a function of some parameter. There are two lines

on each graph, blue for ‘dark’ and red for ‘bright’ conditions. The y -axis is logarithmic, and both graphs show a decreasing trend as the x -axis increases. (c) : This graph shows the wavelength shift and Purcell factor (P.f.) as a function of a scaling parameter α . The wavelength shift (blue line) increases gradually, while the Purcell factor (red line) increases more steeply. (d) : This is the most complex part of the figure, showing radiation patterns and electric field distributions for different scaling parameters (α).

- The top row represents ‘dark’ conditions, while the bottom row represents ‘bright’ conditions.
- There are three columns, each corresponding to a different α value (0.04, 0.3, and 1.5).
- For each condition, there is a 3D representation of the radiation pattern (yellow shapes) and a 2D color map of the electric field distribution.
- The radiation patterns change from a complex shape at $\alpha = 0.04$ to simpler, more symmetric shapes as α increases.
- The electric field distributions also change, becoming more concentrated and intense (shown by red and blue colors) as α increases.

There is a color scale at the bottom indicating ‘nonlocal’ to ‘local’ behavior as α increases from left to right. The figure also includes some additional details:

- A small inset in graph (c) showing the orientation of the nanostructure.
- Axis labels (x, y, z) on some of the 3D plots.
- A scale bar ($p = 3 \mu\text{m}$) on one of the electric field plots.

This figure seems to be demonstrating how the optical properties and behavior of a nanophotonic structure

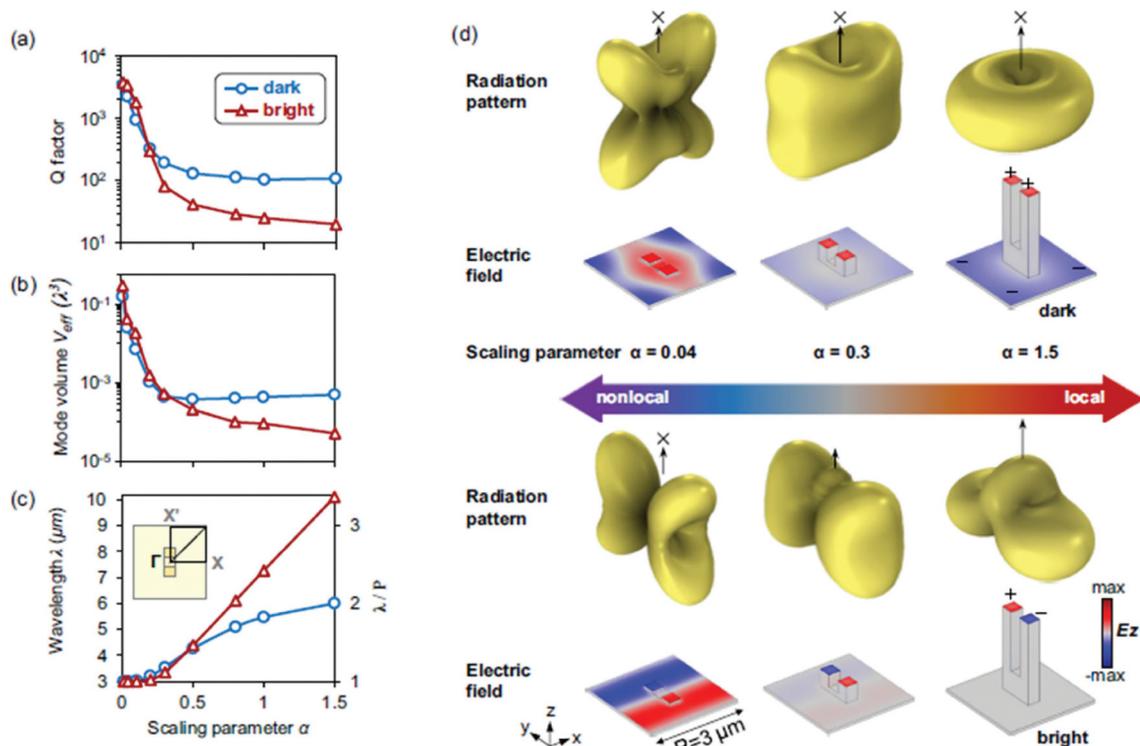


Figure A1. An example of figure involved in genAI retrieval (Section 2.2). This was taken from <https://doi.org/10.1103/PhysRevLett.133.053801> [60] with permission from American Physical Society.

change as it is scaled, possibly transitioning from a nonlocal to a local regime.

In addition to the figure descriptions, a short summary was provided at the beginning. Compared to the original figure caption with less than 100 words, the figure description by the genAI has more than 250 words and tends to explain the contents in a more comprehensive manner. These descriptions were added to the metadata for implementing the RAG in Section 2.2, based on the knowledge of a specific research field, nanophotonics in this study.

Let us show another example of extraction from a figure, presenting a CD effect regarding transmission. A periodic array of hexagonal boron nitride (hBN) nanostructures served as a metasurface. The original figure caption of Figure A2 is as follows [61]:

Linear properties. (a) Schematic image defining different transmission parts. (b) Electrical field distribution in the meta-atom for light with a wavelength of 1494 nm. Theoretical (c) and experimental (d) transmissions of the metasurface for LCP and RCP light. The theoretical CD (ω) is shown in the inset in (c).

Information retrieval from Figure A2 using the genAI described in Section 2.2 yielded the following descriptions:

(a) A schematic image defining different transmission parts of the metasurface. (b) An electrical field distribution in the meta-atom for light with a wavelength of 1494 nm. The image shows a color-coded representation of the electrical field, with purple indicating lower intensity and yellow indicating higher intensity. (c) A theoretical transmission graph of the metasurface for LCP (Left Circularly Polarized) and RCP (Right Circularly Polarized) light. The inset shows the theoretical CD(ω) (Circular Dichroism as a function of frequency). (d) An experimental transmission graph of the metasurface for LCP and RCP light, showing the actual measured results.

In comparison with the original and extracted captions for Figure A2, the descriptions are almost double in the extracted one and more informative than the original. The abbreviations are written down explicitly. In particular, the inset in Figure A2(c) is properly described without missing the information; the simulated CD shows a small value of 0.01 at the maximum, which is in contrast to the ideally perfect CD (i.e. $|CD| = 1$ and $\max(T_{RHC}, T_{LHC}) = 1$) as shown in Figures 5 and 6. Thus, information retrieval from Figure A2 was appropriately implemented.

We furthermore show an example in Figure A3, presenting refractive-index sensing using Si metasurfaces. The original figure caption is as follows [62]:

Measurement results of the refractive index sensitivity. (A), (B) Normalized reflectance spectra measured at different concentrations of IPA for the metasurfaces ($s/L = 2\%$) with and without nanogaps, respectively. Insets show the SEM images with the scale bars of 1 μm . (C) Resonance peak shift of the metasurfaces with (solid line) and without (dashed line) nanogaps as a function of the bulk refractive index. (D) FOM (red open circles) and SNRres (blue closed circles) with different asymmetries s/L for the metasurfaces with nanogaps.

The retrieved caption of Figure A3 is shown below, which is approximately double for the original and more informative. Especially, the main result in Figure A3(c) is explicitly described regarding the sensitivity; thus, the retrieved caption itself is self-contained, needless to consult the text.

A. Normalized reflectance spectra measured in different concentrations of IPA for metasurfaces with nanogaps ($s/L = 2\%$). The graph shows multiple curves representing water and IPA concentrations of 10%, 20%, 30%, and 40%. An inset SEM image shows the metasurface structure with nanogaps. B. Normalized reflectance spectra measured at different concentrations of IPA for the metasurfaces

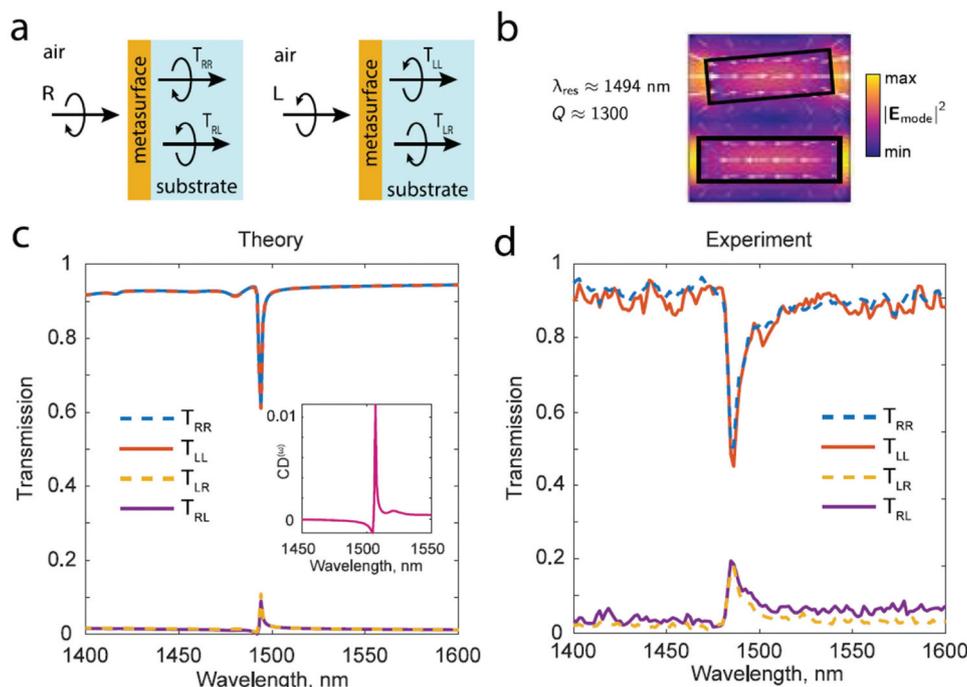


Figure A2. Another example of figure retrieved using the genAI (Section 2.2). Reprinted with permission from [60]. Copyright 2024 American Chemical Society.

without nanogaps. The graph shows similar curves to Figure 6(a) but with different peak positions. An inset SEM image shows the metasurface structure without nanogaps. C. Resonance peak shift of the metasurfaces with (solid line) and without (dashed line) nanogaps as a function of the bulk refractive index. The graph shows that the metasurface with nanogaps has a higher sensitivity (317 nm/RIU) compared to the one without nanogaps (117 nm/RIU). D. Figure of Merit (FOM) and Signal-to-Noise Ratio (SNR_{res}) with different asymmetries s/L for the metasurfaces with nanogaps. The graph shows how these parameters change as the asymmetry (s/L) increases from 1% to 4%.

These retrieved captions were saved in json format, which has low readability for the human eye. To provide a better understanding, let us show one of the examples just below, which is related to Figure A3.

```
{'doi': 'https://doi.org/10.1515/nanoph-2022-0565',
  'authors': ['Keisuke Watanabe', 'Masanobu Iwanaga'],
  'title': 'Nanogap enhancement of the refractive-metric sensitivity ...',
  'journal_name': 'Nanophotonics',
  'volume': 12, 'page_number': 99, 'publication_year': '2023',
  ...
  {'page_number': 6,
   'figure_components': [
    {'figure_name': 'Figure 6a',
     'figure_caption': 'Figure 6: Measurement results of the refractive ...',
     'figure_description': 'Normalized reflectance spectra measured in ...'},
```

```
{'figure_name': 'Figure 6b',
  'figure_caption': '(B) Normalized reflectance spectra measured in ...',
  'figure_description': 'Normalized reflectance spectra measured in ...'},
 {'figure_name': 'Figure 6c',
  'figure_caption': '(C) Resonance peak shift of the metasurfaces with ...',
  'figure_description': 'Resonance peak shift of the metasurfaces with ...'},
 {'figure_name': 'Figure 6d',
  'figure_caption': '(D) FOM (red open circles) and SNRres (blue closed ...',
  'figure_description': 'Figure of Merit (FOM) and Signal-to-Noise Ratio ...'}
]],
...
}
```

Note that we inserted line breaks in the json format, to be comprehensive to readers, and that we omitted parts that are not directly related to Figure A3, which is Figure 6 in reference [60]. We quoted 'figure_description' part as the retrieved caption above, while the 'figure_caption' part is the same with the original caption.

Overall, the information retrieval from figures described in Section 2.2 was implemented in a sufficient manner. We can declare that the preparation of metadata by retrieving the thousands of published papers was implemented in a satisfactory manner.

Appendix B. Examples of Q&A by experts

We show five examples regarding Q&A on nanophotonics. The answers were output by the RAG based only on the metadata (Section 2.2).

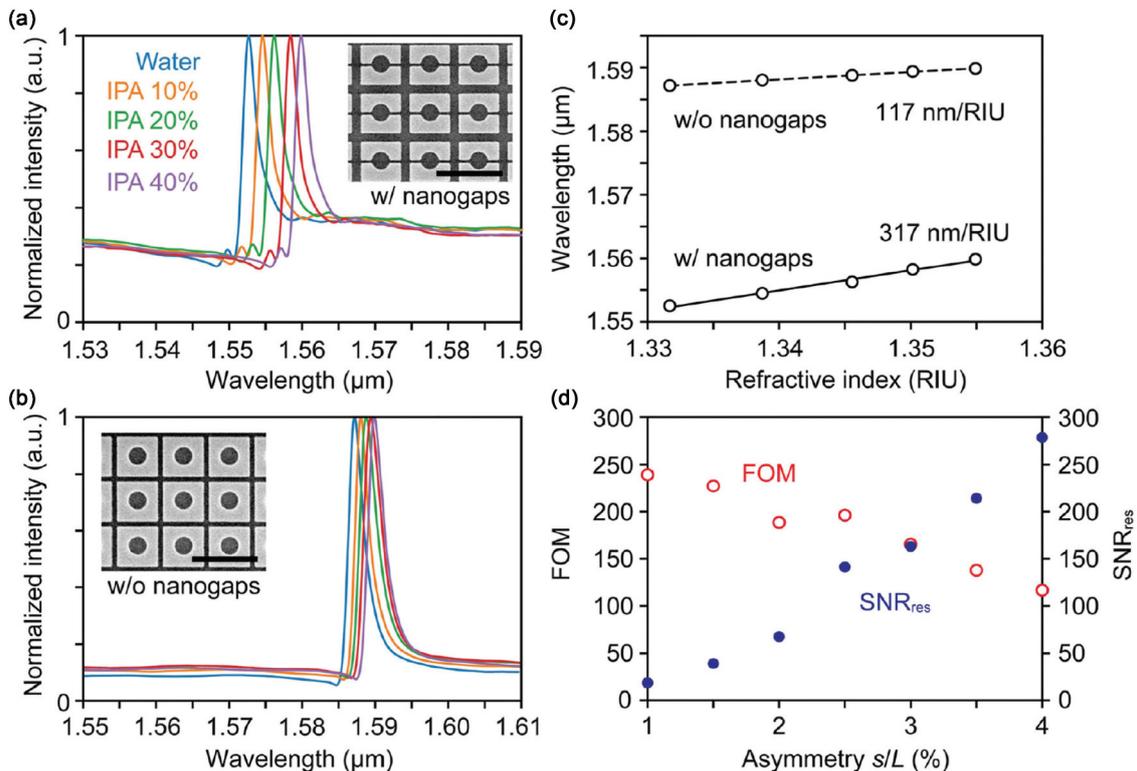


Figure A3. A further example of figure retrieved using the genAI (Section 2.2). Reprinted from [62] under the Creative Commons Attribution 4.0 International License.

Question: Can you tell me the status of optical sensing using metasurfaces?

Answer: Optical sensing using metasurfaces has seen significant advancements in recent years. Metasurfaces, which are two-dimensional arrays of subwavelength resonant nanostructures, offer exceptional capabilities for manipulating light at the nanoscale. This has led to the development of highly sensitive and versatile optical sensors with numerous applications. Some key advancements in metasurface-based optical sensing include:

- (1) Enhanced sensitivity: Metasurfaces can confine light into nanoscale electromagnetic hotspots, greatly enhancing light–matter interactions. This allows for the detection of minute changes in refractive index, enabling highly sensitive chemical and biological sensors.
- (2) Multifunctionality: A single metasurface can be designed to perform multiple functions simultaneously, such as focusing, polarization control, and spectral filtering. This versatility allows for the creation of compact, integrated sensing devices.
- (3) Improved spectroscopy: Metasurfaces have enabled advancements in various spectroscopic techniques, including surface-enhanced Raman spectroscopy (SERS) and surface-enhanced infrared absorption (SEIRA). These improvements allow for more sensitive and selective detection of molecular species.
- (4) Chiral sensing: Metasurfaces with engineered chirality have shown promise in detecting and distinguishing chiral molecules, which is crucial for many biological and pharmaceutical applications [1–3].
- (5) Label-free detection: Many metasurface-based sensors allow for rapid, label-free detection of analytes, simplifying sample preparation and reducing costs.
- (6) Miniaturization: The planar nature of metasurfaces enables the creation of ultra-thin, lightweight optical components. This has led to the development of compact, portable sensing devices suitable for point-of-care diagnostics and field applications.
- (7) Integration with microfluidics: Combining metasurfaces with microfluidic systems has enabled real-time, on-chip sensing capabilities for various biomedical and environmental monitoring applications.
- (8) Tunability: Active metasurfaces, incorporating materials with tunable optical properties, allow for dynamic control of sensor responses and adaptability to different sensing scenarios.

- (9) Quantum sensing: Recent research has explored the integration of metasurfaces with quantum emitters, opening up new possibilities for ultra-sensitive quantum sensing applications.

These advancements have positioned metasurface-based optical sensors at the forefront of sensing technology, with potential applications ranging from medical diagnostics and environmental monitoring to food safety and security. Ongoing research continues to push the boundaries of sensitivity, selectivity, and functionality in these innovative sensing platforms [4–7].

This answer was judged to be correct. Various sensing was referred to with citing related references, which were chosen from our metadata on the AWS. In short, this is a feature of genAI that is good at summing up information. Compared to the short question, the answer is informative. Note that the quoted references above are different from those in this article, which is in common with the following examples.

Question: What are the examples of promising CMOS-compatible materials? It is desired to use nanophotonics materials that can tightly confine light in the visible or infrared regime. Dielectric materials are preferable.

Answer: For visible and infrared light confinement in CMOS-compatible nanophotonics, several promising dielectric materials have been identified: 1. Highly doped silicon 2. Germanium 3. Silicides 4. Germanides. These materials have shown good performance in the technologically important range of 8–12 μm wavelengths, which is relevant for thermal radiation and biological applications. They can achieve propagation lengths of 40 μm with mode sizes of 60 μm at 12 μm wavelength [1].

Additionally, other CMOS-compatible materials that have attracted attention for integrated nanophotonics in the visible spectrum include: 1. Titanium dioxide (TiO_2) 2. Silicon nitride (Si_3N_4) 3. Gallium phosphide (GaP). These transparent materials are being explored to realize compact, high-efficiency optical devices at visible wavelengths. Gallium nitride (GaN), which is highly transparent at visible wavelengths, is also considered promising for active, nonlinear, and quantum nanophotonic applications [2].

It is worth noting that while not explicitly stated as CMOS-compatible, other high-refractive-index dielectrics like silicon (Si) and gallium arsenide (GaAs) have also been widely used in dielectric nanostructures for various nanophotonic applications. These materials have shown potential for applications such as lasing, holograms, sensing, and nonlinear harmonic wave generation [3].

This answer was judged to be correct. The answer seems to be widely collected from metadata, being summing up in a good manner.

Question: What is the shortest operating wavelength achieved in a metasurface?

Answer: The shortest operating wavelength for metasurfaces in the ultraviolet spectrum reported in the given search results is 266 nm. This falls within the deep-UV range and represents a record-short wavelength for metasurface operation. These metasurfaces were implemented using hafnium dioxide (HfO₂), a CMOS-compatible, wide-bandgap, and low-loss dielectric material [1,2]. The achievement of metasurfaces operating at such short wavelengths is significant because it overcomes previous limitations in the UV range. Prior to this, metasurfaces had been demonstrated in the near-UV range down to 355 nm using niobium pentoxide (Nb₂O₅), and in the mid-UV range down to 290 nm using crystalline silicon. However, the efficiency of silicon-based devices was limited due to severe absorption loss at frequencies above silicon's bandgap [3].

This answer was judged to be correct. The question was simple and therefore the RAG answer was easily verified.

Question: Are there any important and unexplored problems in nonlinear optics of metasurfaces?

Answer: While metasurfaces have shown great promise in nonlinear optics, there are still several unexplored areas and challenges in this field:

- (1) Improving conversion efficiency: Although metasurfaces have demonstrated relatively high conversion efficiencies for nonlinear optical processes like second-harmonic generation, there is still room for improvement. Researchers are working on optimizing metasurface designs to further enhance nonlinear interactions and increase conversion rates.
- (2) Broadband nonlinear response: Developing metasurfaces that can maintain strong nonlinear effects across a wide range of frequencies remains a challenge. This would be particularly useful for applications in optical signal processing and broadband optical devices.
- (3) Phase-matching constraints: While metasurfaces can help relax phase-matching constraints in

nonlinear processes, fully overcoming these limitations is still an area of ongoing research [1–3].

This answer was judged to be correct. A future issue regarding nonlinear metasurfaces was asked. The answer was evaluated to be informative and suggestive by an expert.

Question: Can metasurfaces contribute to fluorescence enhancement? If yes, describe the best result obtained to date.

Answer: The highest fluorescence enhancement factor reported for metasurfaces in the given search results is 289-fold. This enhancement was achieved using a silver (Ag) metasurface obtained from a 4 nm-thick mass equivalent layer. The enhancement was observed for a specific concentration (C₃) of Rhodamine 6 G dispersed in PMMA, which was coated on top of the metasurface structures at a deposition rate of 6000 rpm for 40 s [1,2]. It is worth noting that the enhancement factors varied significantly depending on the metal used and the thickness of the deposited layer. For instance, with 2 nm-thick mass equivalent layers, aluminum (Al) metasurfaces provided a 204-fold enhancement, while gold (Au) and silver (Ag) metasurfaces offered 17.3-fold and 102-fold enhancements, respectively. The enhancement factors were also influenced by the geometry and distribution of the metallic nanostructures on the substrate surface [3–5].

This last example was judged to be incorrect. More than 1000-fold fluorescence enhancement was experimentally shown in 10 cases [33]. The RAG failed to pick up the results. Generally, genAI outputs answers based on the 'nearest-neighbor' plausible information they accessed; thus, locally optimal results are possibly output with a certain probability, instead of the genuine, correct results. In this case, frequent Q&A could improve the results, which is often experienced using commercial genAI.

As seen in the five Q&A examples above, the expert questions had large variations and sometimes become vague. The precision of answers by the RAG depends on the questions, which is not limited to this study and is a general property of genAI. When questions are not so clear, the answers tend to be also unclear, containing correct and incorrect contents. This is not a matter of the RAG but that of how to raise questions. When we excluded the ambiguous questions in our trial, the correctness of answers was estimated to be approximately 90%, as noted in [Section 2.2](#).