

Why I Built Image Guard

An Artist's Experience of Digital Image Theft and the Development of RobArt Image Guard

Contents

Before Image Guard: An Artist's Experience of Digital Image Theft	3
Copyright on Paper and Copyright in Practice	11
From Epiphany to Proof of Concept: The Birth of RobArt Image Guard	21
Technical Foundations of Digital Image Protection	28
Invisible Digital Watermarking	31
LSB Watermarking	34
Why Multiple Watermarking Techniques Can Coexist	36
XMP Metadata	37
Cryptographic Hashes and Certified Timestamps	39
Application to Image Guard and a Certificate of Proof	41
Implementation: From Concept to a Built System	43
Conclusion: Returning Control to the Creator	48

Before Image Guard: An Artist's Experience of Digital Image Theft

From Drawing to Digital Vulnerability

My relationship with drawing began long before the internet became part of ordinary creative life. One of the first drawings I can clearly remember making was a copy of an E. H. Shepard illustration of *Winnie-the-Pooh*. My more deliberate creative development, however, began at around the age of thirteen.

Art formed part of the school curriculum at that time, alongside a number of practical subjects including pottery, woodwork and drawing. I enjoyed all of them. They offered something that many of the more academic subjects did not: the opportunity to make something tangible. I was much more comfortable learning through practical activity than through theory alone.

The educational system did not initially encourage that inclination. When I moved into higher school, the academic tier into which I had been placed meant that I was not offered art as a subject. Instead, I was assigned technical drawing. There was certainly creativity involved in technical drawing, but it also contained precisely the theoretical and procedural qualities that had made other subjects difficult for me to enjoy.

One positive influence during this period was a classmate who became a close friend and who would later go on to work professionally in animation, including work connected with Steven Spielberg's animation company. He had been permitted to study art, and I would occasionally accompany him to his lessons simply to watch.

Eventually, the art teacher allowed me to join an after-school class. That small opportunity became significant. It ultimately led to me being permitted to study art at higher levels of my education, where the subject combined both practical work and art history.

At the same time, I was drawing constantly at home. I used whatever media I could find or afford. My parents were not particularly enthusiastic about my interest in art and offered little encouragement, but I persisted.

Failing my A levels on the first attempt was therefore a serious setback. Had my grandmother not stepped in to support me, I would probably have had to leave senior education and find employment. Instead, I was able to continue and eventually entered what was then a Polytechnic to study Media Production.

My drawing ability naturally pushed me towards animation during those years. The course improved my drawing skills considerably, but I eventually realised that I did not particularly enjoy animation itself. I was interested in drawing, not necessarily in making drawings move. I returned increasingly to illustration, and it was from this point that I regard my professional creative career as having begun.

A Career Before the Internet

After completing my studies, I remained for a period in Newcastle, where some of the professional connections surrounding animation were based. It was a pleasant period in many respects, but professionally it did not make me happy.

A subsequent move to Oxford suited me even less, and eventually I returned to Leicester, my home town. There I combined part-time lecturing with illustration work until leaving the United Kingdom in 1999 and moving to Greece.

That move placed me directly into the uncertainty of freelance creative work in a foreign country.

By then, Adobe Illustrator had become an important part of my commercial workflow. Pencil sketches remained the basis of much of what I produced, but the finished work was frequently vector-oriented in appearance. I also made large hand-drawn pieces, although these were generally personal works rather than images created specifically for commercial use.

The early years of freelancing involved a considerable amount of experimentation and difficulty finding work that suited my visual style. Regardless of the commercial direction of individual projects, however, I was always drawing. Drawing was the constant activity underneath everything else.

This period is important to my later experience of copyright because most of my formative years as an artist occurred before the internet became an everyday publishing environment.

Much of my life before higher education had been computer-free. I owned a Commodore Amiga and occasionally used it to experiment with ideas for computer games, but this was recreational rather than central to my artistic work. My first serious encounter with modern computer-based design came during the final year of Polytechnic, when several Apple Macintosh II computers appeared in the design laboratory.

The change was startling.

I therefore belong to a generation of artists whose creative habits were formed in an analogue environment but whose professional lives were gradually transformed by digital tools. I learned to draw before computers became central to design, adopted digital production later, and then witnessed the internet alter not only how creative work was produced and promoted, but also how easily it could be copied.

The Internet as a Creative Tool

My earliest artistic use of the internet was not primarily promotional. It was a reference resource.

Artists have always needed references. If I required a particular pose, an unusual object, an environment or some other visual information, obtaining it directly could be impractical or expensive. I could not simply hire models whenever I required a pose, nor could I obtain every exotic object I might want to draw.

The internet therefore became an extraordinarily useful visual library. I could locate reference material and use it as part of the traditional process of observation, interpretation and drawing.

Its more practical uses were initially connected with employment and business. After moving to Greece, I contacted a local university by email and obtained work within a research organisation, where I became involved with interface design. This drew upon skills from my MA, particularly user-interface analysis and lateral-thinking methods.

The work contained creative elements, but I did not find it especially inspiring. Much of my genuinely creative activity continued to happen in my own time, alongside being a father.

A significant change occurred when I was commissioned to redesign the website of a local radio station. Working on that project forced me to become much more familiar with the technical side of the web. Around the same period I also began using Facebook, which was still comparatively early in its development.

Like many people at the time, I posted work casually.

Completed images and examples of jobs appeared on my personal profile and on a page I created. I uploaded them more or less as they had been produced, often at substantial size. I was thinking about showing the work, not defending it.

That distinction now seems important.

Today, an artist publishing online may instinctively consider resolution, watermarking, metadata, scraping or unauthorised redistribution before uploading an image. At the time, I simply wanted people to see what I had made.

Later I created a Behance portfolio and began placing vector illustrations on stock-art platforms. This followed a particularly productive period in which I had produced design work for Ritzenhoff glassware.

The internet had moved from being a reference resource to becoming part of the infrastructure through which my work could be seen and potentially sold.

It was also becoming the infrastructure through which it could be taken.

Honey and a Change of Direction

Before image theft became a defining issue for me, another event changed the direction of my artwork.

I got a puppy named Honey.

At the time I had been exploring pencil work without any particularly defined objective. One day Honey bounded towards me and I immediately visualised an image. That idea eventually became a drawing called *The Sneaker*.

Something changed after that.

My work became more illustrative and considerably more detailed, and for the first time I began to see a clear opportunity to sell this kind of work online. I created a series of ten illustrations, including works such as *Pack My Trunk*, and uploaded them to numerous print-on-demand sites, online galleries and portfolio platforms.

Again, I uploaded the artwork substantially as it had been created.

The series attracted attention. At one point I was contacted by a prominent online gallery or design blog associated with Adobe and informed that the work had been selected for display.

I was ecstatic.

It felt as though this might be the moment when everything changed: greater visibility, enquiries, followers and potentially substantial new opportunities.

Very little of that materialised.

Something else did.

I began finding copies of *Pack My Trunk* appearing elsewhere on the internet.

They had not been licensed by me.

When Exposure Becomes Appropriation

The unauthorised uses appeared in almost every form imaginable.

My images were copied into scraper blogs. They appeared on social media profiles and as decorative graphical elements. They were reproduced on posters, T-shirts and other products. Some were reposted unchanged, while others were modified.

Two works in particular, *Pack My Trunk* and *Fly on the Wall*, became repeatedly appropriated.

The modification of the artwork was particularly difficult to accept. There is a difference between somebody sharing an image without understanding copyright and somebody deliberately transforming another person's artwork into a new commercial object. Attribution disappeared. Permission was absent. Payment was nonexistent.

In one particularly striking case, I discovered that a company in China had reproduced *Pack My Trunk* as a distressed-effect metal sign.

I investigated the available information and found claims indicating production quantities of approximately one thousand units per day.

Whether every one of those units represented a completed sale was almost beside the point. My artwork had apparently entered an industrial manufacturing process entirely without me.

I received no licence fee.

I received no royalty.

I was not consulted.

Requests for the use to stop went unanswered.

This represented one of the clearest lessons of digital distribution: the person who creates an image and the people capable of monetising it can become completely disconnected from one another.

Copyright theoretically gives the creator ownership, but the internet can give possession of a usable copy to almost anyone.

The difference between those two realities began to dominate my thinking.

Creative Licence and the Attempt to Fight Back

I have always preferred to fight rather than simply accept something I believe is wrong.

As the scale of the infringements became clearer, I became increasingly interested in copyright and enforcement. Eventually I created a website called **Creative Licence**.

It was intended both as an educational resource and as a practical action site.

I researched takedown procedures and created a dedicated takedown form. I assembled information explaining copyright issues, practical guides, resources and procedures that artists could use when confronting unauthorised copying. When I encountered stories in the news about artists experiencing similar problems, I contacted them to exchange information and experiences.

My assumption was that other artists would be strongly motivated to defend their work.

The site never became the success I expected.

The reasons remain difficult to identify. Perhaps many artists feel powerless in the face of large-scale copying. Perhaps the effort required to pursue infringements outweighs the likely reward. Perhaps the psychological burden of repeatedly confronting people who have taken one's work is itself enough to discourage enforcement.

Creative Licence has since been archived.

During the same period, I discovered Pixsy and increasingly used reverse-image-searching tools, including Google's reverse image search.

Initially these technologies seemed like the answer.

They could reveal where an image had travelled.

Unfortunately, they also revealed just how far it had travelled.

Each search uncovered another unauthorised use, followed by another and another.

The technology worked. That was precisely the problem.

Instead of providing reassurance, reverse-image search frequently produced evidence that the situation was much worse than I had known.

When I later began creating caricatures, the problem intensified. At times it appeared that individual images generated more infringements than legitimate sales.

The tools created to help an artist regain control could therefore have the opposite psychological effect: they could expose the almost uncontrollable scale of the loss.

A Sisyphean Form of Enforcement

I continued working and continued sending takedown notices.

Occasionally they succeeded.

The problem was that every success seemed to expose ten additional infringements.

The process became Sisyphean. A single image could be removed from one website while copies continued circulating elsewhere. Another seller could upload it. Another scraper could index it. Another social account could repost it.

There was no final victory.

The fundamental nature of digital copying makes retrospective protection extraordinarily difficult. Once an unrestricted copy of an artwork has circulated sufficiently widely, the creator would effectively need to locate every surviving digital instance and remove it at its source.

That is impossible in practice.

Much of my earlier work is therefore, in my own terminology, **in the wind**.

Copies already exist beyond my control. They may be sitting on websites I have never seen, on private computers, in product databases, in social-media archives or in scraper collections waiting to be republished.

Eventually I largely stopped looking.

There is a point at which searching for infringements stops protecting the artist and begins damaging the artist.

I decided instead to concentrate on what remained good about making art.

Even so, the experience changed my behaviour.

I have produced substantially less personal artwork during the last two years, and I believe the accumulated experience surrounding theft has contributed to the depression I continue to associate with that period.

When I do consider placing commercial artwork online now, I am far more defensive. I use heavy watermarks. I crop images. I avoid allowing viewers to see an unrestricted full version whenever possible.

I no longer upload new work casually to sales platforms because, from my perspective, such environments can also become highly efficient resources for automated scrapers.

My own WordPress site has changed accordingly. I removed its shop, which also reduced the large numbers of automated attempts to access the site. I installed protections intended to discourage simple downloading techniques such as right-click saving.

None of these measures is absolute.

A determined thief with sufficient technical knowledge can usually find another route.

That recognition eventually became central to my thinking.

When Copyright Becomes Someone Else's Message

One infringement demonstrated that image theft is not always principally about lost sales.

I had created a caricature of Joe Biden that I considered successful and followed it with a caricature of Kamala Harris while the two were campaigning together.

I later received a message from somebody who followed my work asking whether I knew that I had appeared on the news in the United States.

I did not.

They sent me the report.

My caricatures had been taken and incorporated into a large roadside political billboard in Maryland. The images had been altered and placed into an aggressively anti-Biden and anti-Harris political message.

The use was entirely unauthorised.

For me, the significance went far beyond copyright.

I am politically socialist in outlook, and the original artwork had been created partly as an expression and celebration of my own political sympathies. The stolen artwork had effectively been turned against the beliefs that contributed to its creation.

An image can therefore be stolen in more than one sense.

Its economic value can be appropriated, but its meaning can also be appropriated.

Someone else had made my work appear to say something I had never intended it to say.

A lawyer contacted me and initially began pursuing the infringement. I believed that copyright protection existed automatically when an artwork was created, and that had always been the basis on which I operated.

The practical reality of attempting enforcement in another jurisdiction proved far more complicated.

Once it became apparent that the artwork had not been formally registered in the United States, the legal action effectively ceased to be commercially attractive and the case was dropped.

I continued trying to have the billboard removed for months.

Later, with assistance from a sympathetic podcaster and radio host, I discovered that the imagery had apparently been used a second time.

Even identifying precisely who was responsible proved difficult.

The experience reinforced one of the most important distinctions I had begun to understand: **possessing a right and possessing an effective means of enforcing that right are not necessarily the same thing.**

From Protection to Identification

For some time my software-development interests were focused almost entirely on Adobe Photoshop.

I had approximately eighteen plugin concepts under development at various stages. I viewed locally installed productivity tools as a comparatively safe and useful area in which to combine creativity and software development.

The first ideas that eventually contributed to image protection did not, however, begin with copyright.

Some years earlier I had watched a documentary concerning covert intelligence operations. It mentioned the possibility of placing information inside image files in ways that were not immediately visible.

The idea remained with me.

Why, I wondered, could similar principles not be applied publicly?

I subsequently discovered academic research exploring methods of embedding information within digital images, including research associated with Oxford University. At that stage it remained primarily an interesting technical question.

Years later, while working in Photoshop, I happened to examine the File Info panel more closely.

I was already familiar with EXIF metadata from digital photography, but the extensibility of XMP metadata suddenly became much more interesting. Photoshop did not merely display information. It allowed information to be added.

A question formed:

Could useful information be written into an image so that it travelled with that image?

Initially, this was not a copyright-protection question at all.

I was interested in embedding notes within image files: information that could remain associated with a document and later be opened or recovered.

That line of thought led to the development of tools including **Noteli** and **JotFlagger**.

Software development itself became increasingly compelling. At roughly the same period I allowed my Gravity Forms subscription to lapse, including the system I had been using around DMCA-related forms. That brought copyright back into focus.

The two areas began to converge.

The result was an idea called **SylentID**.

SylentID attempted to place substantial ownership information within XMP metadata so that an image contained evidence of its provenance. The concept was attractive because the identification existed inside the file rather than merely alongside it on a website.

Its weakness was equally clear.

Metadata can be stripped.

A thief who understands how a file is constructed can remove information deliberately, while many ordinary image-processing operations may discard metadata without malicious intent.

I therefore began combining XMP identification with invisible watermarking. The invisible watermark could act as the primary identifier, while the metadata provided an additional evidential layer.

This was a considerable improvement, but the underlying problem remained.

Once a thief possesses the image, protection becomes a contest between the protections embedded in the file and the technical ability of the person attempting to remove them.

The creator remains reactive.

From Photoshop to WordPress

For a long time I had thought of plugin development almost exclusively in terms of Photoshop.

Then, almost casually, I began wondering whether WordPress plugins were equally accessible to an independent developer.

I asked the question and discovered that they were.

In some respects, particularly visually and in terms of integration with web interfaces, I found WordPress plugin development easier to explore.

I began creating a range of plugins. Many were concerned with system functionality and productivity, although projects such as **Aegra** and **StatFlo** brought me increasingly close to questions of monitoring and security.

I was not consciously trying to create an image-protection system.

Then, recently, while taking a nap, an idea arrived almost completely uninvited.

I had been working simultaneously on a growing family of plugins and on the website through which they would eventually be presented. Copyright protection was not the immediate problem I had been trying to solve.

Nevertheless, two previously separate pieces of thought suddenly connected.

Years of experiencing image theft had shown me the limitations of trying to recover artwork after it had escaped.

My experiments with metadata and invisible watermarking had shown me both the possibilities and weaknesses of embedding evidence inside files.

WordPress development had shown me that I could intervene not only in the artwork itself, but in the mechanism through which a website delivered that artwork.

The question therefore changed.

Instead of asking only how an image could identify itself after it had been stolen, I began asking whether the website could fundamentally change **what was being given away in the first place**.

That was the point at which the path towards **RobArt Image Guard** began.

Copyright on Paper and Copyright in Practice

My experiences with online infringement gradually forced me to recognise an important distinction between **owning copyright and being practically capable of enforcing it**.

At first sight, copyright law appears unusually favourable to the creator. Unlike patents, trademarks and many registered design rights, copyright generally does not begin with an application to a government authority. The international foundation for this principle is the Berne Convention for the Protection of Literary and Artistic Works. Among its central principles are *national treatment*—under which a qualifying foreign author should receive broadly the protection that a country gives its own authors—and *automatic protection*, meaning that copyright protection cannot be made dependent upon registration, deposit, the use of a copyright symbol or some other formality.

For an artist, the principle seems reassuringly simple: create an original protected work and copyright arises. The difficulty begins when that right has to be exercised against somebody who has ignored it.

The European Union: Automatic Rights

Across the European Union, original artistic works receive copyright protection automatically. There is no general EU requirement for an artist to register a drawing, illustration or photograph before copyright can exist. For most authorial works, the harmonised general term of economic protection is the author's lifetime plus seventy years.

EU legislation also gives authors important exclusive economic rights relevant to online image theft. These include control over reproduction of a protected work and its communication or making available to the public. The Enforcement Directive additionally requires Member States to provide civil measures and remedies for intellectual-property infringement.

This does not mean that copyright law is identical throughout Europe. EU legislation has harmonised substantial areas of economic copyright, but national legal traditions still matter, particularly in matters such as authors' moral rights, procedure, evidence and the practical route through which an infringement is pursued.

France

France provides a useful example because its concept of *droit d'auteur* places considerable emphasis upon the relationship between an author and a work.

An original work is protected from its creation without the author having to complete an administrative registration procedure. French law distinguishes broadly between economic rights and moral rights.

The moral-rights element is particularly striking in relation to experiences such as the alteration of an artist's imagery. Article L121-1 of the French Intellectual Property Code gives an author rights relating to respect for their name, status and work and describes this moral right as **perpetual, inalienable and imprescriptible**. Economic exploitation rights normally continue through the author's lifetime and for seventy years following death.

The distinction is important. Copyright infringement does not necessarily concern money alone. Removing an artist's identity, altering a work or placing that work into a context contrary to its creator's intentions can raise questions about the personal connection between creator and creation as well as unauthorised economic exploitation.

Germany

German copyright follows a similarly creator-centred principle. The German Patent and Trade Mark Office explains that copyright comes directly into existence through personal intellectual creation and does not depend upon entry in a register or even upon the presence of a copyright notice.

German law expressly protects the author's personal relationship with the work as well as its economic exploitation. It provides a right to recognition of authorship and a right to object to distortion or other interference with a work where that treatment is capable of harming the author's legitimate intellectual or personal interests. Copyright normally expires seventy years after the author's death.

Again, the principle is broader than simply preventing somebody from making an unauthorised print. The identity of the creator and the integrity of the work itself are recognised as interests deserving protection.

The United Kingdom

The United Kingdom provides another useful comparison. UK copyright is automatic and there is no official government register of copyright works. An artist does not have to apply, pay a fee or register a drawing before copyright can arise.

UK law also recognises moral rights, including rights concerning identification as author and objection to certain derogatory treatments of a work. Derogatory treatment may include distortion, mutilation or another alteration prejudicial to the author's honour or reputation, although the application and formal requirements of individual moral rights vary.

There is also an attempt to make lower-value enforcement more accessible. In England and Wales, copyright claims valued at no more than £10,000 can, where suitable, be taken through the Intellectual Property Enterprise Court's Small Claims Track, and a claimant does not necessarily require a lawyer to use that procedure.

This is significant because one of the persistent problems for individual creators is proportionality. An infringement may be genuine while the amount recoverable is insufficient to justify the cost of conventional litigation.

The United States: Copyright Exists, but Registration Matters

The United States presents perhaps the most important contrast for my own experience.

U.S. copyright does **not** simply begin when a work is registered with the Library of Congress. The U.S. Copyright Office itself states that copyright exists from the moment a qualifying work is created. Registration does not manufacture the underlying copyright.

Registration can, however, radically alter the practical value of enforcement.

For a U.S. work, registration or refusal of registration is generally required before an infringement action can be filed. A qualifying non-U.S. work is treated differently: the U.S. Copyright Office states that a foreign work does not need to have been registered before its owner files an infringement lawsuit in the United States.

This sounds as though the foreign artist is therefore fully protected.

The reality is more complicated.

Under section 412 of the U.S. Copyright Act, timely registration is normally required for access to two particularly important remedies: **statutory damages and attorney's fees**. Registration before infringement begins, or generally within three months after first publication, can preserve eligibility for those remedies.

Where those conditions have not been met, the copyright itself does not disappear. U.S. law still provides for recovery of actual damages suffered by the copyright owner and profits of the infringer attributable to the infringement.

But this produces an important practical difference.

Statutory damages can remove some of the difficulty of proving precisely how much an artist lost because of an infringement. U.S. law currently permits ordinary statutory awards ranging from \$750 to \$30,000 per work in appropriate cases, with a discretionary maximum of \$150,000 where willful infringement is proved.

Without access to statutory damages and potentially recoverable attorney's fees, a claimant may instead have to prove actual financial loss or trace profits attributable to the infringement. A case may therefore remain legally valid while becoming commercially unattractive to pursue.

That distinction retrospectively explains much of my confusion surrounding the unauthorised American political billboard that incorporated my caricatures. My belief that I owned copyright automatically was not fundamentally mistaken. The more consequential issue was the difference between **having copyright and having access to remedies substantial enough to make enforcement economically practical**.

For an independent artist discovering an infringement in another country, that distinction can be devastating.

Metadata and the Law

There is another aspect of modern copyright law particularly relevant to digital-image protection: information identifying the work and its owner.

U.S. law contains specific provisions protecting certain forms of **copyright management information**. Section 1202 of the Copyright Act prohibits specified intentional removal or alteration of copyright-management information, or distribution of works knowing such information has been removed or altered, where the required knowledge relating to infringement exists. Copyright-management information can include the title of the work and identifying information about its author or copyright owner.

European law contains a related concept. Article 7 of the EU Information Society Directive requires legal protection against specified unauthorised removal or alteration of electronic rights-management information and certain subsequent distribution or communication of affected works where the necessary knowledge is present.

This is particularly relevant to the development of digital identifiers, embedded ownership information and technologies such as invisible watermarking. Such mechanisms do not themselves create copyright, and their presence cannot guarantee successful enforcement, but they can contribute to the problem of establishing provenance: **who created a file, what work it represents, and what information accompanied it before unauthorised manipulation occurred**.

From Berne to the DMCA: Automatic Copyright, Digital Enforcement and the Orphan Works Debate

Understanding the present position of an online visual artist in the United States requires looking at several stages in the development of American copyright law. These stages are sometimes compressed into the idea that the United States adopted "automatic copyright" in 1989. The history is more complicated.

The principal modern framework is the **Copyright Act of 1976**, signed in 1976 and largely effective from January 1, 1978. Among its major changes, the Act extended federal copyright protection to both published and unpublished works once an original work was fixed in a tangible medium. In contemporary terms, the Copyright Office continues to describe copyright as arising automatically when an original work is fixed.

This meant that copyright did not have to be created by registration.

However, important formalities still remained, particularly for published works. Most significantly for visual creators, copyright notice continued to have legal importance.

The 1989 Berne Change

The next major transition came through the **Berne Convention Implementation Act of 1988**. The United States formally joined the Berne Convention on March 1, 1989. From that date, copyright notice became optional for works first published on or after March 1, 1989.

This distinction matters.

The 1989 change did not suddenly invent the principle that an artist owned a work upon creating it. Rather, it removed one of the remaining formal requirements capable of causing serious problems for a published work.

Before this transition, failure to comply with notice requirements could, depending upon the circumstances and period involved, have significant consequences for copyright protection. After the Berne implementation, a modern artist could publish a drawing without placing “© Artist Name” visibly upon it and would not lose copyright merely because that notice was absent. The Copyright Office nevertheless continues to recommend the use of copyright notices because they identify ownership and can have evidential and remedial advantages.

For an international artist, the United States joining Berne was especially significant. One of the fundamental ideas of the Berne system is that protection should not depend upon creators completing formalities in every country in which their work might subsequently appear.

This is close to the understanding under which I operated.

I was an English artist living in Europe. I created an original drawing and understood that copyright existed automatically. In that fundamental respect, my understanding was correct.

What I did not initially appreciate was that **automatic ownership and maximum practical enforcement are different things**.

The United States retained registration as an enormously important procedural and remedial tool even though registration was not what created the underlying copyright.

That contradiction—or at least what can feel like a contradiction to a foreign artist—would later become painfully relevant to my own experience.

Copyright Meets the Internet: The DMCA

Less than a decade after the United States joined Berne, widespread internet use created an entirely different copyright environment.

The **Digital Millennium Copyright Act of 1998**, or DMCA, became one of the principal American legislative responses to copyright in the digital environment. It added several mechanisms that are particularly relevant to online artists.

Three parts are especially relevant to the history described in this paper.

Section 512: Notice and Takedown

Section 512 established a system of conditional “safe harbors” for qualifying online service providers. In simplified terms, providers that meet the statutory requirements can receive limitations on copyright liability for material placed on their services by users.

One component of this framework became what artists commonly call the **DMCA takedown**.

A copyright owner who discovers infringing material hosted through a qualifying provider can submit a notice identifying the copyrighted work and the allegedly infringing material. The service provider’s

compliance with the statutory framework forms part of the conditions associated with its safe-harbor protection. The system also includes a counter-notification procedure for users who believe material was removed incorrectly.

For an independent artist, this was potentially powerful.

It created a mechanism through which infringement could sometimes be addressed without immediately filing a federal lawsuit.

It was also the system that informed much of my later work with takedown procedures and ultimately contributed to the creation of Creative Licence.

Yet it contains the same fundamental limitation that I repeatedly encountered:

| the takedown process begins after the image has already escaped.

The artist must discover the copy, identify where it is hosted, locate the appropriate service provider, prepare the notice and then wait for action.

If the same image has been copied to twenty other locations, the process begins again twenty times.

The DMCA therefore provided an enforcement mechanism for online infringement, but it did not remove the underlying scalability problem that had made digital image theft so damaging.

The internet could reproduce an image considerably faster than its creator could issue notices concerning those reproductions.

Technological Protection Measures

Section 1201 of the DMCA addressed another aspect of digital copyright: **circumvention of technological protection measures**.

The legislation prohibits certain forms of circumvention of technological measures controlling access to copyrighted works and also restricts specified technologies or services primarily associated with circumvention. The provisions are complex and contain exemptions, including exemptions considered through a recurring Copyright Office rulemaking process.

The significance for the history of Image Guard is conceptual rather than a claim that Image Guard automatically receives protection under section 1201.

By 1998, American copyright law was explicitly recognising that digital copyright involved more than simply declaring ownership of the underlying artwork.

| The technology controlling access to the work could itself become part of the legal framework.

This is a significant development when considering modern image protection.

A traditional copyright statement says:

You may not reproduce this image without permission.

A technological protection system attempts to add another proposition:

I will also make unauthorised acquisition or access more difficult in the first place.

The distinction would eventually become central to my own thinking.

Section 1202: Copyright Management Information

The DMCA also created a provision particularly relevant to metadata and digital identification.

Section 1202 protects certain forms of **Copyright Management Information**, commonly abbreviated to CMI.

CMI can include information identifying a work, its author or copyright owner, and other specified information associated with rights and permissions. The statute prohibits particular forms of knowingly false CMI and certain intentional removal or alteration of CMI where the statutory knowledge requirements relating to infringement are satisfied.

It is important not to simplify this into the statement that *removing metadata is always illegal*. Section 1202 contains specific requirements concerning conduct, knowledge and intent, and whether a particular metadata field, identifier or technical system qualifies as protected CMI can depend upon the circumstances.

Nevertheless, the principle is extremely relevant.

Digital identifying information is not merely a technical convenience. American copyright law recognises that information connecting a creative work with its author, owner and rights can itself have legal importance.

This provides an interesting background to my later experiments with XMP metadata, SylentID and invisible identifiers.

Those experiments began largely from the practical question:

| Can an image carry evidence of who it belongs to?

The DMCA demonstrates that the legal system had independently recognised the importance of much the same relationship: maintaining reliable information connecting a digital work to the people and rights associated with it.

The Unexpected Consequence of Automatic Copyright

Automatic copyright solved an obvious problem for creators.

Artists should not lose ownership simply because they failed to complete a bureaucratic procedure.

However, eliminating formalities created another problem.

When registration and copyright notices are optional, there is no guarantee that every protected work has a corresponding public ownership record.

An image may therefore remain fully copyrighted even when the person wishing to use it cannot determine who owns it.

This is the origin of the modern **orphan works** problem.

The Copyright Office defines the issue broadly around copyrighted works whose owners cannot be identified or located. It began a substantial investigation into the problem in 2005 and issued a major report to Congress in 2006.

An orphan work is therefore not the same thing as a work in the public domain.

Nor does a work cease to be copyrighted merely because its creator cannot be found.

The work may still possess a perfectly valid copyright.

The problem is that a prospective user cannot locate the person from whom permission must be obtained.

Why Digital Images Are Particularly Vulnerable to Becoming “Orphaned”

For visual artists, there is a particularly uncomfortable dimension to this problem.

Images detach from their origins extraordinarily easily.

A photograph or illustration may begin on the artist's own website accompanied by:

a name;

copyright notice;

title;

contact details;

XMP or IPTC metadata;

licensing information;

a link to the artist's portfolio.

Someone copies the image.

It is uploaded elsewhere.

Another person downloads that version.

A social-media system recompresses it.

A marketplace produces a derivative JPEG.

Metadata disappears.

The filename changes.

The artist's signature is cropped away.

Eventually somebody encounters the image with no information identifying its creator.

The copyright has not vanished.

| The identity has.

The Copyright Office has itself recognised the particular problem of visual works losing identifying information, noting that the removal of such information makes it harder both for creators to trace uses and for legitimate prospective users to locate the owner, thereby contributing to orphan-work problems.

This creates a disturbing paradox for a digital artist.

An image can effectively be made harder to attribute **because other people have already copied and redistributed it.**

The more widely an unauthorised copy spreads, the greater the possibility that future copies become separated from information identifying the original creator.

A living artist can therefore discover an image described informally as "owner unknown" while being very much alive, commercially active and entirely capable of licensing it.

An orphan work is not necessarily the work of an absent creator.

Sometimes it is a work from which the creator's identity has been removed.

The Proposed Orphan Works Legislation

The U.S. Copyright Office concluded in its 2006 investigation that the orphan-works problem was genuine and recommended a legislative response.

Congress subsequently considered several proposals.

These included the **Orphan Works Act of 2006**, followed by the **Shawn Bentley Orphan Works Act of 2008** in the Senate and a separate **Orphan Works Act of 2008** in the House.

The Senate bill, S.2913, actually passed the Senate by unanimous consent on September 26, 2008. It did not, however, complete the legislative process and did not become law. Congress as a whole did not enact the proposed general orphan-works regime.

This point is essential because discussions of orphan works online have sometimes created the impression that American law now allows people to appropriate any image whose creator cannot immediately be identified.

It does not.

Finding an unattributed picture on the internet does not automatically transform that picture into an orphan work that anyone is free to use.

The principal proposals considered a much narrower system.

A user would have been required to undertake a good-faith or reasonably diligent search for the copyright owner before relying upon the proposed limitation of remedies. If the owner subsequently emerged, the proposals generally contemplated reasonable compensation while restricting some of the remedies that would otherwise have been available to the copyright holder.

The Copyright Office described the approach as a limitation upon remedies rather than destruction of the copyright itself. The owner did not cease to own the work merely because another person had been unable to find them.

That distinction is fundamental.

The proposal was not:

“We cannot find you, therefore your copyright no longer exists.”

It was closer to:

“We made the required search but could not find you; if you later appear, the legal consequences of our use may be limited.”

For many visual artists, however, that distinction offered limited comfort.

Why Visual Artists Objected

Photographers, illustrators and graphic artists raised particularly strong concerns during the orphan-works debate.

Their problem was structural.

An author publishing a novel normally has a title page, publisher, ISBN and extensive textual information through which ownership may eventually be traced.

A digital image may consist of nothing more than several million coloured pixels.

Once identifying information has been separated from those pixels, searching for its creator can be considerably more difficult.

Visual-art organisations therefore argued that orphan-works legislation could disproportionately affect creators whose works routinely circulate independently of identifying context. Concerns included the cost and burden of maintaining searchable records for very large bodies of work and the danger that a legitimately copyrighted image might become practically difficult to identify after redistribution.

The 2008 proposals attempted to address some of these issues, including special considerations involving searchable databases for visual works. Nevertheless, the debate exposed a fundamental conflict.

Libraries, archives, historians and subsequent creators may possess works they would legitimately like to reproduce but for which no owner can realistically be located.

At the same time, an illustrator may look at exactly the same system and ask:

What happens when the reason you cannot find me is that somebody else removed my name from my image?

That is not a theoretical concern in my experience.

I have repeatedly discovered my artwork copied without attribution, altered, reposted and converted into products.

Each reproduction potentially creates another version of the work increasingly disconnected from me.

Orphaned in Appearance, Not in Ownership

For digital artists, I believe it is useful to distinguish between a work that is genuinely historically orphaned and one that has become **orphaned in appearance**.

The first might be an old photograph discovered in an archive for which extensive research genuinely fails to reveal the photographer or their successors.

The second might be an illustration created by a living artist five years ago that has been copied through twenty websites until every identifying element has disappeared.

From the perspective of the twentieth website, both may appear similar:

creator unknown.

Legally and ethically, their histories could hardly be more different.

This distinction reinforces the value of provenance systems.

Visible attribution can be removed.

Metadata can be stripped.

A filename can be changed.

A website containing ownership information can disappear.

The more independent methods by which an artwork can remain connected with its creator, the harder it becomes for that relationship to vanish accidentally—or deliberately.

This is one reason the evolution from XMP metadata towards invisible identifiers became increasingly interesting to me.

It was not simply about catching somebody who had copied an image.

It was also about preventing the image from losing its identity.

A Strange Circle in Copyright History

There is an irony in the development of American copyright law.

The removal of formalities was intended to protect creators.

An artist should not lose decades of copyright because a notice was omitted from a print.

The United States therefore progressively moved towards automatic protection, culminating in its Berne-compliant removal of mandatory notice in 1989.

But removing those formalities also meant that future users could encounter copyrighted works for which no authoritative ownership record was readily available.

The orphan-works debate was partly an attempt to address that consequence.

For the visual artist, this produces a difficult balance.

I strongly support the principle that my artwork belongs to me from the moment I create it. I should not have to ask a government office for permission to own my own drawing.

My experiences in the United States have nevertheless taught me that **optional registration, optional copyright notices and removable digital metadata can leave an enormous gap between theoretical ownership and practical identification.**

The lesson is not that automatic copyright was a mistake.

The lesson is that automatic copyright requires effective ways of maintaining the connection between the work and its creator.

That problem becomes more severe, not less, in an internet environment where copies can become detached from their original source within seconds.

The DMCA attempted to address parts of this new environment through takedowns, technological-protection provisions and protection for copyright-management information.

The orphan-works debate exposed another part of the same problem.

All ultimately revolve around one question:

How does a digital work retain a meaningful connection to the person who created it once that work begins travelling through systems the creator no longer controls?

That question lies directly behind my experiments with metadata, invisible watermarking and eventually RobArt Image Guard.

The Enforcement Gap

Looking across these jurisdictions reveals a recurring pattern.

The law generally begins from a remarkably strong proposition: the artist creates an original work and receives rights without first asking permission from the state. International copyright rules deliberately reject registration as a condition for the existence of those fundamental rights.

Yet infringement happens in a different world.

A copied image may be hosted in another jurisdiction. The operator may be anonymous. A marketplace may disappear and reopen under another name. The infringer may have little money even if identified. Legal representation may cost more than the likely recovery. Evidence may have to establish when a work was created, where it was first published, what was copied, who copied it and what financial damage resulted. In some jurisdictions, procedural actions such as timely registration can dramatically affect the remedies available even though they do not determine whether copyright originally existed.

This is the gap between **copyright as a legal right and copyright as a practical defence.**

It is also the point at which my interest began to move away from relying exclusively upon enforcement after infringement.

Takedown notices, reverse-image searches, lawyers and litigation all begin with essentially the same unfortunate fact:

the image has already been taken.

The question that eventually led towards RobArt Image Guard was different.

Rather than concentrating solely upon what an artist can do *after* an unrestricted digital image has escaped into the internet, I began to wonder what could be done **before that unrestricted image was ever handed over in the first place.**

From Epiphany to Proof of Concept: The Birth of RobArt Image Guard

The development of RobArt Image Guard did not begin with a conventional software specification. It began with a question that appeared unexpectedly while I was resting.

There was, however, a longer intellectual background to that moment.

During my Master's studies in lateral thinking, I encountered an anecdote attributed to Albert Einstein suggesting that difficult problems could be carried into a period of rest, allowing ideas to develop as conscious, logical control diminished. I subsequently adopted a version of this practice, deliberately allowing creative or technical problems to remain in mind while drifting towards sleep. The precise Einstein attribution appears difficult to substantiate and may partly reflect later accounts of similar techniques associated more reliably with Thomas Edison and Salvador Dalí, both of whom are reported to have used the moment of falling asleep — holding an object that would drop and wake them at the edge of unconsciousness — to capture ideas arising in that state.

The underlying phenomenon, however, is no longer merely anecdotal. The transition into sleep passes through a brief hypnagogic phase, the first stage of non-REM sleep known as N1, and controlled experiments have identified it as an unusually productive state for creative insight. Lacaux, Andrillon, Bastuji, Oudiette and colleagues, reporting in *Science Advances* in 2021, recreated a version of the Edison technique in the laboratory and found that participants who briefly entered N1 before being woken were substantially more likely to discover a hidden rule governing a problem than those who had remained awake or fallen into deeper sleep — a result the authors described as a creative “sweet spot” at sleep onset.

A related line of work goes further, attempting not merely to observe the effect but to direct it. Researchers at the MIT Media Lab, including Horowitz, Cunningham, Maes and colleagues, developed a method they term *targeted dream incubation*, in which a chosen theme is introduced to a sleeper during the hypnagogic period and its influence on subsequent thought is measured. Their experiments indicate that deliberately seeding a specific subject at sleep onset can measurably shape and enhance later creative performance on tasks related to that subject.

The practice I had adopted intuitively, in other words, corresponds closely to a phenomenon that has since received experimental support.

(Note for review: the two research citations above — Lacaux et al., Science Advances, 2021, on the N1 creative sweet spot; and the MIT Media Lab targeted dream incubation work — should be confirmed against the original sources, including exact author lists and publication details, before release.)

The idea that eventually became RobArt Image Guard emerged during one of these periods, although on this occasion I had not deliberately set out to solve the problem of image theft.

I had been thinking about software development.

From Photoshop Plugins to WordPress

At that point, most of my plugin development had been centred upon Adobe Photoshop.

I had approximately eighteen Photoshop plugin ideas or projects in various stages of development. I had begun to regard local productivity software as a comparatively safe creative direction. Instead of relying entirely upon selling digital artwork in an environment where that artwork could be copied, I could use my creative experience to build tools for other artists and designers.

One of these projects was SylentID.

SylentID had grown out of my earlier experiments with XMP metadata and invisible identification. Its purpose was to associate an image with substantial ownership information that could travel within the file itself. Invisible watermarking could provide one identification mechanism while XMP metadata could provide another.

At around the same period, I began wondering whether WordPress plugins were as accessible to independent development as Photoshop plugins.

After establishing that they were, I started experimenting.

I quickly found WordPress development attractive, particularly because many interface and visual elements could be created comparatively directly. A growing range of plugin ideas followed. Many were concerned primarily with productivity or site-management functions, although projects such as Aegra and StatFlo had already taken me towards monitoring, system behaviour and aspects of security.

RobArt Image Guard was not part of this planned range.

It effectively interrupted it.

The Placeholder Question

While resting, I had been thinking about SylentID and the possibility of eventually releasing it through the Adobe marketplace.

The idea of covert tagging and embedded information was therefore already somewhere in my thoughts.

Then, for reasons I cannot completely reconstruct, I began thinking about placeholder images in web development.

A placeholder image normally has little intrinsic value. It exists to occupy space, establish dimensions, indicate where content will eventually appear or provide a temporary visual substitute.

This produced an unexpectedly simple question:

| What is the real value of a placeholder image to an image thief?

The answer appeared to be: almost none.

A scraper wants the photograph, illustration or commercial artwork. It does not want a blank or meaningless substitute.

That observation caused two previously separate ideas to connect.

If an ordinary website gives the browser the valuable image directly, then protection mechanisms are largely placed *around something that has already been delivered*. Right-click restrictions can inconvenience the least technical visitor. Metadata can identify the file. Watermarks can discourage reuse. Invisible identifiers can potentially survive after an image has escaped.

But the valuable image has still been placed into the public delivery system.

What if it were not?

The idea began developing quickly.

Instead of placing the original image within the ordinary public WordPress uploads structure, the original could be placed somewhere private.

A placeholder could occupy its normal public position.

The two could then be associated through a unique identifier.

Something within the website could recognise that identifier and, when the legitimate page was viewed, cause the intended image to appear in place of the placeholder.

To WordPress and to ordinary inspection, the public asset would remain the placeholder.

The original would have no conventional public URL.

In its earliest form, the concept could therefore be summarised as:

| **public placeholder, private original, permanently paired.**

The next question was whether such a system could actually work.

Externalising the Idea

After waking, I began discussing the concept with ChatGPT.

The first step was not to write a finished plugin but to formalise the idea into a design document. The purpose was to determine whether the architecture was technically plausible and identify the major components that would be required.

I subsequently shared that design with Claude, which at that stage was my principal development environment for the WordPress implementation.

Claude agreed with much of the underlying architecture.

Rather than immediately attempt to create a complete commercial plugin, we decided to build a **Proof of Concept**, or PoC.

This distinction was important.

The PoC did not need to solve every problem associated with image protection. It needed to answer one much narrower question:

| **Could the fundamental placeholder/private-image substitution mechanism actually work inside WordPress?**

The First Proof of Concept

The first implementation was deliberately simple.

Not every image on a WordPress website requires protection. Interface graphics, icons, decorative elements and other non-sensitive assets may gain nothing from being processed through a protection system.

Image Guard therefore needed to distinguish between ordinary and protected uploads.

The initial PoC intercepted the normal WordPress **Add Media** process and introduced a modal choice allowing the user to decide whether an image should be uploaded conventionally or treated as protected content.

For a protected upload, the underlying idea was an upload swap.

The valuable source image would be separated from the normal public media structure and placed into private storage. A standard placeholder of the required dimensions would then represent that image within WordPress.

For this initial experiment there was no need for sophisticated placeholder generation, watermark editors or extensive configuration. A default placeholder of an appropriate size was sufficient.

The objective was simply to establish that WordPress could work with one image while the visitor ultimately saw another.

After refining the requirements, the first complete test worked almost immediately.

A protected image was uploaded.

The WordPress Media Library displayed the placeholder.

The placeholder was inserted into a Page as though it were an ordinary WordPress image.

The Page was then viewed on the live site.

The placeholder disappeared from the visible result.

The original image appeared.

The basic proposition conceived while resting had become demonstrable software.

A Small but Important Demonstration

One particular behaviour made the success immediately tangible.

When the displayed image was right-clicked, the browser's normal image-saving behaviour referred to the placeholder.

Choosing **Save Image As...** therefore saved the placeholder rather than the artwork that appeared to be visible on the page.

This was not, by itself, the security breakthrough.

Preventing right-click saving has existed for many years and can be circumvented easily by knowledgeable users. Image Guard was not intended to present disabling a context menu as meaningful security.

What mattered was what the behaviour demonstrated.

The image understood by the public WordPress structure and the image perceived by the visitor had successfully become **different resources**.

The Media Library knew about the placeholder.

The Page contained the placeholder.

The normal save operation encountered the placeholder.

Yet the viewer saw the intended protected image.

The conceptual separation had worked.

What the PoC Did — and Did Not — Prove

Success immediately led to a more difficult question.

Could the image still be obtained?

The answer was yes.

This limitation did not originate specifically in Image Guard. It originated in the fundamental operation of a web browser.

For a browser to display an image, sufficient image data must eventually reach the browser to allow that image to be rendered. A technically knowledgeable person with access to browser developer tools can examine network activity and page behaviour at a substantially deeper level than a normal visitor.

Image Guard could not repeal that architectural reality.

This forced an important decision about the philosophy of the project.

The objective would **not** be to claim that digital images could be made impossible to acquire.

Such a claim would be technically unrealistic.

Instead, the objective would be to remove the easiest and most scalable methods of acquiring them.

The early development assessment was that the PoC might already obstruct a very substantial proportion of ordinary image theft. Internally, we discussed figures in the region of 80 percent for user-driven acquisition and potentially close to complete disruption of the simple automated scraping methods the system had specifically been designed to defeat.

These figures were not empirical measurements and should not be interpreted as statistically established protection rates. They represented our initial assessment of the categories of theft addressed by the design.

What mattered was the shift in effort.

A visitor who ordinarily right-clicked and saved an image would receive the placeholder.

A simple scraper reading page markup would encounter the placeholder.

A crawler enumerating publicly accessible WordPress uploads would encounter placeholders and ordinary public assets, but not the privately held protected originals.

The remaining route required a person with sufficient technical knowledge to inspect what the browser was doing and retrieve displayed content individually.

That distinction was considered acceptable.

Changing the Economics of Theft

The decision reflected an increasingly important principle behind Image Guard.

Security does not always require making an action theoretically impossible.

Sometimes it is sufficient to make the action sufficiently inconvenient that its economics change.

Much of the image theft I had personally experienced depended upon scale.

Automated systems can collect enormous quantities of imagery because the marginal effort required to obtain the next file is almost zero. A scraper does not become tired. It does not have to open each image manually, inspect developer tools, identify the correct resource and save files individually.

It can simply enumerate predictable resources.

This is precisely the behaviour Image Guard was intended to disrupt.

If a scraper collecting ten thousand images instead collects ten thousand useless placeholders, the process has failed from the scraper's perspective.

A technically knowledgeable individual might still be able to recover a specific displayed image.

But having to do this manually, deliberately and individually is a fundamentally different proposition from bulk acquisition.

The aim therefore became not absolute prevention but **friction**.

For casual theft, substantial friction.

For automated theft, potentially prohibitive friction.

For the most determined technically capable user, sufficient friction that mass acquisition becomes increasingly impractical.

Defence in Layers

Accepting that a sufficiently determined person could still obtain a displayed image led naturally back to the technologies I had already been investigating.

Image Guard did not need to depend upon a single protective mechanism.

The placeholder/private-original architecture could form the first layer.

Additional protections could then be offered when required.

These included:

visible watermarking;

invisible watermarking;

XMP ownership and provenance information.

Importantly, these were not intended to form a mandatory processing pipeline.

Different publishers have different priorities.

A photographer may be willing to place a prominent visible watermark across a preview image. An artist presenting a portfolio may regard that as visually destructive. A commercial site may want invisible identification but no obvious mark. Another site may simply want protection from automated scraping.

Image Guard therefore began developing around the principle of **optional layered protection**.

The central protected-delivery architecture would remain the foundation.

The owner could then decide whether additional measures were appropriate.

In conceptual terms, the system had become:

**private protected image → publicly represented placeholder → controlled substitution on the live page
→ optional visible watermark → optional invisible watermark → optional or partially automated XMP
identification.**

Each component addressed a different problem.

The placeholder architecture addressed acquisition.

Visible watermarking addressed deterrence and unauthorised presentation.

Invisible watermarking addressed later identification.

XMP information addressed provenance and ownership information associated with the digital file.

None was expected to solve the entire problem independently.

Together, however, they represented a substantially different approach from conventional image-protection plugins.

A Different Definition of Success

The development of Image Guard also forced me to reconsider what successful digital-image protection actually means.

An absolute standard would make the project impossible.

If success were defined as:

“No human being must ever be capable of obtaining a copy of an image they can see on a screen,”

then practically every browser-based protection system must fail.

A screenshot alone demonstrates the fundamental difficulty.

A more useful definition is:

Can the valuable source file be removed from routine public access, can automated acquisition be frustrated, can casual theft be substantially reduced, and can additional evidence or deterrence be attached to images where stronger protection is required?

That is the problem Image Guard is attempting to solve.

The distinction is important because security software can easily become misleading when it promises an absolute outcome that the underlying technology cannot provide.

Image Guard was instead beginning from an explicit acknowledgement of its boundary.

There would remain a technically sophisticated route through which a determined individual could acquire what the browser ultimately rendered.

We decided to accept that boundary rather than distort the system around an impossible objective.

The remaining question was how much harder everything else could be made.

The End of the Proof-of-Concept Stage

At the time of writing, RobArt Image Guard is only in its second day of active development.

The architecture described here should therefore not be mistaken for documentation of a completed commercial product.

The Proof of Concept has demonstrated the central proposition:

a WordPress site can publicly represent a protected image through a placeholder while holding the associated original outside normal public access and presenting the intended image to the legitimate visitor through a separate delivery mechanism.

That result was enough to justify continuing.

The next stage is considerably larger.

A practical plugin will require decisions concerning user experience, image sizing, WordPress compatibility, caching, responsive image handling, watermarking, metadata, performance, security, failure states, interaction with themes and page builders, and numerous behaviours that a deliberately minimal PoC was never intended to address.

The earlier draft of this account was written on the second day of development, and it stopped at that point because those stages had not yet been attempted.

They have since been built.

What began as a thought occurring somewhere between wakefulness and sleep has now moved well beyond its first encounter with working software. The sections that follow the technical foundations describe what was actually implemented, and — more importantly for a paper of this kind — the reasoning behind the specific choices made along the way. Whether the result becomes a robust, practical and broadly usable method of protecting creative work on the web is still a question for time and use to answer, but it is no longer a question of whether the architecture can exist. It exists.

Technical Foundations of Digital Image Protection

The development of RobArt Image Guard draws upon several technologies that existed long before the plugin itself was conceived.

These technologies solve different problems.

A **visible watermark** attempts to discourage unauthorised use by placing ownership information directly into the appearance of the image.

An **invisible digital watermark** attempts to place recoverable identifying information within the image data without materially changing what the viewer sees.

XMP metadata stores structured information about the image, its creator, rights and history inside or alongside the file.

A **cryptographic timestamp** can provide independent evidence that a particular digital object or cryptographic representation of it existed at a particular time.

None of these technologies should be regarded as absolute proof or absolute protection. Their value comes from combining deterrence, identification, provenance and evidence so that defeating one layer does not necessarily defeat all of them.

Visible Watermarking

From Paper to Pixels

The term *watermark* considerably predates computers.

For centuries, papermakers incorporated identifying marks into sheets during manufacture. Watermarks could identify a mill, manufacturer or source while remaining part of the material itself. The European tradition is particularly associated with centres such as Fabriano, whose papermaking and watermark history extends over seven centuries. UNESCO notes that historic European paper mills used distinctive watermarks capable of identifying their manufacture.

The modern photographic or artistic watermark adopts the same underlying idea but implements it visually.

A name, logo, copyright symbol, website address or other identifying graphic is superimposed upon an image before that image is distributed.

Unlike an invisible digital watermark, the purpose is specifically to be seen.

What a Visible Watermark Is Intended to Do

A visible watermark has several functions.

First, it identifies or advertises the source of the image.

Second, it reminds the viewer that the image is controlled rather than simply an unclaimed file found online.

Third, it reduces the commercial usefulness of an unauthorised copy.

An image copied from a website may technically remain usable, but a large ownership mark passing through important image detail can make that copy unsuitable for a poster, product, advertisement or clean republication without further work.

This is the basic deterrent model used by commercial image libraries.

Getty Images, for example, continues to present watermarked preview imagery while distinguishing those previews from the unwatermarked content available through licensing.

The watermark therefore performs a useful economic function.

It does not make copying impossible.

It makes the copied version **less valuable**.

Why Some Visible Watermarks Are More Effective Than Others

A watermark placed discreetly in an image corner is visually attractive but comparatively weak.

A thief can often crop it away.

Similarly, a small mark placed over a smooth sky, blank wall or other low-detail region may be relatively straightforward to reconstruct.

A more defensive watermark intersects meaningful visual information.

This produces the familiar compromise between protection and presentation.

The stronger the watermark, the more difficult the unauthorised restoration may become; yet the same watermark may also interfere with the legitimate viewer's enjoyment of the artwork.

Research into visible watermarking consequently treats visibility, security and preservation of image quality as competing requirements.

In practical terms, a strong visible watermark may use several characteristics:

placement through important rather than disposable image regions;

sufficient size that simple cropping is ineffective;

semi-transparency allowing the underlying artwork to remain visible;

variation in luminance across the watermark so that it interacts with both light and dark areas;

repetition or multiple marks where appropriate;

identification that actually means something, such as an artist name, site or rights reference.

The objective is not merely to put a logo on the picture.

It is to make the watermark sufficiently integrated with the image that removing it cleanly becomes more expensive in time and effort than simply obtaining a properly licensed image.

The Getty Example

Getty Images is often mentioned informally as an example of effective visible watermarking.

That should be interpreted carefully.

It would be difficult to support an academic claim that a Getty watermark is technically impossible, or even uniquely difficult, to remove. Modern image-processing and generative inpainting techniques have made automatic watermark detection and reconstruction an active research area.

The more useful lesson from Getty is strategic.

The preview copy and the licensed copy serve different purposes.

The preview is sufficient for evaluation, but visibly identifies itself as a preview. The customer receives the clean version through the licensing process. Getty's current systems continue to make this distinction explicitly.

This principle has direct relevance to Image Guard.

A site owner should be able to decide that some publicly delivered derivatives require visible deterrence while the valuable private master remains clean.

Limitations of Visible Watermarking

Visible watermarking is deterrence, not cryptographic security.

Watermarks can be cropped.

They can be cloned over manually.

They can sometimes be removed by reconstruction or generative image-processing systems.

If a mark occupies only a small or visually simple region, restoration may be comparatively easy.

A sufficiently aggressive watermark can make removal much more difficult, but there is ultimately no perfect balance because the legitimate viewer must still be able to see the artwork.

For Image Guard, visible watermarking therefore makes most sense as an **optional additional protection layer**, not as the foundation of the system.

Invisible Digital Watermarking

Visible watermarking changes what the viewer sees.

Invisible watermarking attempts to change what the image **contains**.

The goal is to embed information within an image in a form that remains imperceptible, or nearly imperceptible, during ordinary viewing but can subsequently be detected or decoded.

Digital watermarking became an established research field during the early 1990s. Van Schyndel, Tirkel and Osborne published their influential work *A Digital Watermark* in 1994, examining the feasibility of encoding information within digital images.

By the later 1990s, research had moved strongly toward making these hidden signals resistant to ordinary image processing and deliberate attack. Cox, Kilian, Leighton and Shamoon's widely cited 1997 work on spread-spectrum watermarking proposed inserting an imperceptible signal into significant spectral components of multimedia and examined resistance to compression, filtering, cropping, scaling and other transformations.

This introduces an important distinction.

A hidden mark is not necessarily a **robust** mark.

It is easy to hide information in an image.

The difficult problem is making that information survive what happens to the image afterwards.

DCT Watermarking

What DCT Means

DCT stands for **Discrete Cosine Transform**.

The mathematical transform itself was introduced by Nasir Ahmed, T. Natarajan and K. R. Rao in 1974.

Rather than treating an image simply as individual pixel values, the DCT represents image information in terms of spatial-frequency components.

This has become particularly important in image compression. DCT-based processing forms part of traditional JPEG compression, where image blocks are represented through transform coefficients rather than stored simply as uncompressed pixels.

That relationship makes DCT especially interesting for watermarking.

How DCT Watermarking Works

A simplified DCT watermarking process can be described as follows:

the image, or blocks within the image, are transformed from the spatial domain into DCT frequency coefficients;

selected coefficients are modified according to the information that is to be embedded;

the inverse DCT converts the modified frequency representation back into a normal-looking image;

a detector later examines the relevant coefficients to determine whether the expected watermark signal remains.

The information being embedded does not need to be a visible logo.

It may be a short numerical identifier, ownership code, verification token or other payload.

Many systems deliberately embed into selected middle-frequency regions. Very low-frequency components strongly influence visible image structure, so excessive alteration can become perceptible. Very high-frequency information, by contrast, is more likely to be discarded by compression and filtering. Modern DCT watermarking research therefore frequently selects coefficients intended to balance imperceptibility against robustness.

Why DCT Is Attractive for Copyright Protection

A well-designed transform-domain watermark can survive operations that would completely destroy simple attached metadata.

The image may be renamed.

It may be copied onto another website.

It may be re-saved.

It may undergo moderate compression.

Certain DCT-based schemes have demonstrated robustness against operations including JPEG compression, filtering and noise, although effectiveness depends heavily upon the specific algorithm, embedding strength and attack.

The important feature is that the identifying signal forms part of the image information itself.

It is no longer merely a text field accompanying the image.

This makes DCT watermarking useful for **ownership detection after separation from the original website or metadata**.

Making a DCT Watermark Effective

The strength of a DCT watermark involves several competing factors.

Embedding strength matters. A signal that is too weak may disappear after compression. A signal that is too strong may visibly degrade the artwork.

Coefficient selection matters. Choosing components likely to survive anticipated transformations improves robustness.

Redundancy matters. Repeating or distributing the encoded information can allow recovery even where part of an image has been damaged or cropped.

Secret-key selection can matter. Research systems frequently use keys to determine embedding positions or transform the payload so that simply knowing the broad watermarking method is insufficient to reproduce or remove the mark reliably.

Error correction can also improve survival. Research has combined DCT embedding with error-correcting codes so that damaged payload bits may still be reconstructed.

For an application such as Image Guard, the embedded payload does not necessarily need to contain an artist's complete name and contact information.

A short robust identifier can be more useful.

That identifier can point to a protected database record containing the richer information.

Limitations of DCT Watermarking

DCT watermarking is robust only in relative terms.

Heavy transformation can destroy the signal.

Aggressive cropping may remove enough embedded regions to prevent detection.

Rotation, geometric distortion and resizing can interfere with synchronisation unless the detector compensates for them.

Severe compression can destroy selected coefficients.

Modern generative reconstruction introduces an even more fundamental problem because an image can be semantically recreated rather than merely recompressed.

No invisible watermarking method should therefore be described as indestructible.

The more appropriate objective is **survivability under the transformations that commonly occur during ordinary theft and redistribution.**

LSB Watermarking

What LSB Means

LSB stands for **Least Significant Bit**.

Digital pixel values are represented numerically. In an 8-bit channel, for example, a value can range from 0 to 255.

The least significant bit contributes only the smallest possible numerical change to that value.

Classic LSB embedding exploits this fact.

If a pixel value is represented in binary, altering its final bit changes the numerical value by only one. Such a change is normally visually imperceptible.

LSB therefore provides a comparatively straightforward method of hiding binary information directly within image pixels. It is one of the established spatial-domain approaches to watermarking and data hiding.

How LSB Embedding Works

Suppose an 8-bit pixel value is:

10000110

which represents decimal 134.

If the information bit that needs to be hidden is 1, the final bit can be changed:

10000111

The value becomes 135.

The visual difference is generally negligible, yet one binary bit of information has been stored.

Repeating the process across many pixels can store a considerably larger payload.

Research implementations have used one or more LSB positions, selected pixels, secret keys and various distribution techniques to improve capacity or resistance to simple attacks.

Advantages of LSB

LSB has several attractive properties.

It is computationally simple.

It can hold substantial amounts of hidden information.

Changes can be extremely difficult to perceive visually.

It can therefore be useful where processing speed or payload capacity matters.

It can also be useful for **fragile watermarking**, where the aim is not necessarily to survive alteration but to reveal that alteration has occurred. Research has used LSB-based approaches specifically for authentication and tamper detection.

The Main Weakness of LSB

The same property that makes LSB visually unobtrusive makes it vulnerable.

The embedded data occupies precisely the portion of the pixel representation that image processing can alter most freely without producing obvious visual damage.

Lossy compression, resizing, filtering, colour conversion or deliberate bit manipulation can therefore destroy simple LSB payloads.

For that reason, classic LSB schemes are generally much less appropriate than robust transform-domain systems when the objective is to maintain an ownership identifier after uncontrolled internet redistribution.

This does not make LSB useless.

It gives it a different role.

A fragile LSB layer can indicate whether the file remains substantially unchanged, while a more robust frequency-domain mark attempts to survive modification. Hybrid research has investigated exactly this combination of spatial and frequency-domain techniques.

For Image Guard, such distinctions are important because a watermark does not need to perform only one task.

One signal might identify ownership.

Another might indicate tampering.

A third might identify the particular protected asset or delivery instance.

Why Multiple Watermarking Techniques Can Coexist

There is no requirement that an image contain only one watermark.

Different layers can have different purposes.

A possible layered model might therefore contain:

Visible watermark

Deters casual reuse and identifies the source immediately.

Robust DCT watermark

Attempts to retain an invisible ownership or asset identifier after compression and redistribution.

Fragile or semi-fragile watermark

Helps identify alteration or establish whether a particular file remains substantially identical to the protected version.

This follows the broader idea of defence in depth.

One layer does not need to be perfect if an attacker must discover and defeat several unrelated layers independently.

The important requirement is that the protections do not collectively damage the artwork to the point that protection defeats the purpose of publishing it.

XMP Metadata

What XMP Is

XMP stands for **Extensible Metadata Platform**.

Adobe announced XMP as a metadata initiative in 2001. It was designed to provide a common framework for creating, processing and exchanging metadata across digital publishing workflows.

XMP subsequently became standardised through ISO 16684.

Adobe's current specification describes XMP as a standardised data model and serialization system that can be embedded within formats including JPEG and PDF while allowing software that does not understand XMP to continue reading the underlying file.

This is fundamentally different from digital watermarking.

A watermark modifies the image signal.

XMP is structured **information about the file**.

What XMP Can Contain

XMP can contain information such as:

creator;

title;

description;

copyright statement;

rights information;

creation and modification dates;

keywords;

source information;

document identifiers;

custom application-specific properties.

Adobe's XMP namespaces include Dublin Core properties such as dc:creator, dc:title and dc:rights, together with rights-management and document-identification properties.

This makes XMP particularly attractive for copyright management.

Instead of embedding only a small numerical identifier, a file can carry a comparatively rich ownership record.

For example:

Creator: Artist name

Copyright: © Artist, year

Rights: All rights reserved

Website: creator's site

Asset ID: unique Image Guard identifier

Certificate ID: proof certificate reference

Creation Date: recorded date

Watermark ID: identifier recoverable from invisible watermark

XMP can also support custom namespaces, allowing a system such as Image Guard to create structured fields specifically for its own provenance model.

Why XMP Is Useful for Verification

XMP is highly useful because the information is machine-readable.

Software can extract it automatically.

A verification tool can compare an embedded Image Guard asset identifier with a database entry.

A file received years later could potentially reveal the creator, copyright information, internal certificate reference and other provenance data without needing to know the filename or original website location.

This directly addresses one of the problems discussed earlier in relation to orphan works.

The more ownership information that accompanies an image, the easier it may be to reconnect that image with its creator.

Adobe specifically identifies author and copyright information as appropriate uses for embedded XMP metadata.

The Weakness of XMP

XMP has one enormous limitation.

It is removable.

An application may deliberately strip metadata.

A social-media platform may discard it while processing an upload.

An image optimizer may create a derivative containing little or no metadata.

A thief can deliberately remove the entire XMP block while leaving the visible image substantially unchanged.

XMP should therefore never be treated as the sole ownership mechanism.

Its strength is richness.

Its weakness is persistence.

This is precisely why combining XMP with invisible image watermarking is attractive.

The XMP record can contain extensive ownership and provenance information while the invisible watermark contains a smaller identifier capable of reconnecting the image with that record if the metadata is removed.

Cryptographic Hashes and Certified Timestamps

Watermarking and metadata address the question:

| Can this image identify itself?

A certified timestamp addresses a different question:

| Can we establish that this particular digital object existed by a particular time?

This is potentially extremely important for copyright evidence.

Cryptographic Hashing

A cryptographic hash function converts digital data into a fixed-length value.

If even a small change is made to the underlying file, its cryptographic hash will ordinarily change.

The hash can therefore act as a digital fingerprint of a specific sequence of file data.

Instead of attempting to register an entire high-resolution image with an external service, a system can calculate a strong cryptographic hash of that image and submit the hash.

The original artwork does not necessarily need to leave the creator's system.

RFC 3161 Trusted Timestamping

One established standard for this process is the **RFC 3161 Time-Stamp Protocol**, published in 2001.

The standard describes communication with a **Time Stamping Authority**, or TSA.

RFC 3161 explicitly describes timestamping as providing evidence that a particular datum existed before a particular time.

The process is conceptually straightforward.

A cryptographic hash of the data is calculated.

That hash, or *message imprint*, is submitted to the TSA.

The TSA associates it with a trustworthy time and produces a cryptographically signed timestamp token.

Importantly, RFC 3161 specifies that the TSA timestamps the **hash representation**, rather than requiring inspection of the original data itself.

Later, the file can be hashed again.

If the new hash matches the value represented by the timestamp token and the TSA's signature validates correctly, this provides cryptographic evidence that the corresponding data existed no later than the timestamped moment.

What a Timestamp Does Not Prove

This distinction is essential for an academic discussion of copyright.

A timestamp does **not automatically prove authorship**.

If I steal somebody else's photograph today and timestamp it tomorrow, the timestamp proves that I possessed that particular data by tomorrow.

It does not prove that I took the photograph.

Likewise, a timestamp alone does not create copyright.

Its value lies in supporting a larger chain of evidence.

If an artist possesses:

the original high-resolution master;

development or source files;

consistent XMP creator information;

an invisible watermark identifier;

a database registration record;

a cryptographic hash;

and an independently certified timestamp created near the time of publication or registration;

the evidential position becomes considerably stronger than relying upon a filename or manually entered creation date alone.

Qualified Electronic Timestamps in Europe

The European eIDAS framework gives electronic timestamps explicit evidential recognition.

Article 41 provides that an electronic timestamp cannot be denied legal effect or admissibility as evidence merely because it is electronic, while a **qualified electronic timestamp** receives a presumption concerning the accuracy of its date and time and the integrity of the data to which the timestamp is bound.

This makes the distinction between an ordinary computer clock and a recognised timestamping authority particularly important.

A date written into an XMP field can be changed.

A file-system creation date can change when a file is copied.

A date printed on a PDF certificate is simply a statement.

A cryptographically verified timestamp from an independent authority represents a fundamentally different category of evidence.

Application to Image Guard and a Certificate of Proof

This is where the various technologies can begin to converge.

A future Image Guard **Certificate of Proof** need not merely be a decorative PDF claiming that an image belongs to somebody.

It could represent a verifiable evidence package.

At the moment an image is registered or protected, Image Guard could create an evidence record containing information such as:

Image Guard asset ID;

creator or copyright owner;

original filename;

MIME type;

pixel dimensions;

file size;

cryptographic hash of the private master;

relevant EXIF data;

relevant XMP data;

visible-watermark status;

invisible-watermark identifier;

plugin or processing version;

copyright-registration information where supplied;

certificate identifier.

A cryptographic hash of the master — and potentially a canonical evidence manifest associated with it — could then receive an RFC 3161 timestamp from an independent TSA.

The resulting timestamp token could be retained with the private evidence record and referenced from the Certificate of Proof.

The certificate could consequently make a much more defensible statement:

At the certified time stated on this record, a digital file producing this cryptographic fingerprint existed and was associated by the Image Guard system with this evidence record.

That is a much stronger proposition than:

This PDF says I made this picture on Tuesday.

It still would not allow Image Guard itself to declare legal ownership conclusively.

Courts, registration authorities and competing evidence ultimately determine disputed rights.

What Image Guard could do is preserve a **coherent chain of technical evidence**.

The private master establishes the underlying file.

The cryptographic hash identifies that exact file state.

The trusted timestamp establishes proof of existence by a given time.

XMP records structured ownership and provenance information.

The invisible watermark gives distributed copies a possible route back to the asset record.

The visible watermark discourages straightforward unauthorised commercial reuse.

The Certificate of Proof presents those separate pieces in a human-readable form.

This suggests an important principle for the later development of RobArt Image Guard:

the strongest protection may not come from attempting to make a digital image impossible to steal, but from combining barriers to acquisition with persistent identification and independently verifiable evidence of provenance.

That would transform Image Guard from a simple anti-download mechanism into something broader.

It would become a system concerned with three stages of digital copyright protection:

Prevention — make acquisition harder.

Identification — allow a separated image to retain or recover a connection to its owner.

Evidence — preserve verifiable information capable of supporting that ownership claim later.

No single one of these stages is complete by itself.

Together, they represent a substantially more realistic model for protecting digital creative work.

Implementation: From Concept to a Built System

The sections above were, in their original form, an account of an idea and the technologies behind it. Much of the language was necessarily provisional. A protection layer *could* work in a particular way; a certificate *would* eventually contain particular fields. This final section describes what was actually built, and concentrates less on the fact that it was built than on the reasoning behind the specific decisions taken. In a paper of this nature the *why* is more useful than the *what*, because the choices reveal where the honest boundaries of the system lie.

Throughout the implementation, one principle was treated as non-negotiable: the system should never claim more than it can demonstrate. Digital image protection is an area unusually prone to overstatement, and a tool that promises the impossible discredits itself the moment a knowledgeable user tests it. Every decision described below was made with that constraint in mind.

Protected delivery, and what the browser inspector actually reveals

The foundation remained the placeholder architecture proved in the proof of concept: the valuable master is held in private storage outside the normal public uploads structure, a placeholder occupies the public position, and the intended image is substituted only when a legitimate page is viewed. Two decisions shaped how far this could be pushed.

The first was to make the substitution deliver a *resolution-limited* rendering rather than the master itself. The public never needs the full-resolution original in order to see the picture on a page, so the delivered image is capped to the size the page actually displays. This means the strongest copy anyone can obtain through ordinary viewing is a display-sized derivative, not the master. The master genuinely does not enter the public delivery path at all.

The second decision was to make the delivered image reach the browser through a signed, single-use request rather than a predictable public URL, and to hand it to the page as a temporary in-memory object rather than a saveable file. The practical consequence of this is worth stating precisely, because it is easy to overstate. During development we tested the most common technical route a determined user is assumed to take: opening the browser's developer tools, examining the network activity and the page source, and attempting to open or save the underlying resource. The placeholder that appears in the public files is a small, low-value image. The request that delivers the real rendering has no stable, re-openable address, and the rendered image is held as a temporary object that does not exist as a file the browser will save. Attempting to open that object in a new tab, or to save it from the inspector, produced nothing useful.

This is a genuine and reproducible result, and it matters because the developer-tools route is exactly the method most people assume defeats any such protection. But it must be framed honestly. It does not make the displayed pixels impossible to capture. A screenshot of a screen is always possible, and no browser-based system can prevent it. What the result demonstrates is narrower and still valuable: the master never leaves the server, the public can only ever obtain a display-capped copy, and the usual "just grab it from the inspector" shortcut does not yield the file. The honest promise is therefore not *unstealable pixels*. It is that the original is withheld, casual and automated acquisition is defeated, and the copy that can be captured is limited, watermarked and traceable.

Visible watermarking as independent layers

The visible watermark was built as three independent layers — text, logo and signature — each with its own position, size, opacity, rotation, tiling and fine offset, and each able to be applied on its own or together.

This design was not chosen for the sake of generality. It came from a specific, real use pattern. An artist protecting a piece frequently wants two quite different marks on the same image at once: a large, faint

ownership wordmark spread across the picture as the anti-theft deterrent, and a small, discreet copyright line along an edge for attribution. Those two marks share almost nothing — different size, different opacity, different placement, different intent. A single watermark control, however flexible, forces them to compromise. Treating each mark as an independent layer with its own complete set of controls removes the compromise entirely: because no settings are shared between layers, there is nothing for them to clash over. The signature layer extends the same idea to a hand-drawn mark, captured on a drawing pad so that a genuine signature, rather than a typed name, can travel with the image.

The watermark is composited into the delivered derivative, not the master. The master is never altered by the visible mark, which means a watermark can be changed or removed and simply re-composited on the next request. This preserves a clean original while still ensuring the public only ever sees the marked version.

Where the invisible marks are written, and why in the browser

The forensic (invisible) mark and the embedded copyright metadata are written into the stored master rather than the delivered derivative. This is a deliberate reversal of the visible-watermark decision, and the reason is provenance. A visible watermark is a public deterrent and belongs on the public copy; an invisible forensic mark and an ownership record are evidence, and evidence is only useful if it travels with the file that might one day be found elsewhere. Writing them into the master means that if the master ever escapes, the identifying information goes with it.

The forensic mark's embedding was implemented to run in the site administrator's own browser rather than on the server. The reason is practical rather than architectural preference: the frequency-domain techniques that make an invisible mark robust, and the error-correction that helps it survive redistribution, are well served by mature, permissively licensed libraries that run in the browser, whereas performing the same work reliably on an arbitrary shared web host cannot be assumed. Running the embedding in the browser, using the master fetched privately for that purpose, keeps the server's workload negligible and avoids depending on capabilities a typical host may not provide. After embedding, the marked image is read back immediately and the recovered payload confirmed, so the operation reports success only when the mark is demonstrably present and recoverable — there is no silent claim that a mark was applied.

Honest cryptography for embedded copyright data

The copyright metadata is written as a standard, embeddable record inside the image file, following the same principle as earlier metadata experiments but with the cryptography made genuinely meaningful.

Earlier work in this lineage protected its embedded data with a reversible obfuscation — a simple transformation combined with encoding. That approach hides data from casual inspection and lets a specific application recognise its own records, which is useful, but it is not security in any strong sense, because the transformation can be reversed by anyone who examines it. Presenting it as encryption would overstate what it does. The implementation therefore replaced obfuscation with real, standard cryptography, and offers three clearly distinguished modes rather than a single ambiguous one.

The first mode writes the copyright record in the ordinary, readable form that any standards-compliant tool can display. Its purpose is interoperability and public attribution: the copyright line, creator and rights statement appear in any image application's file-information panel.

The second mode writes the same readable record but attaches a keyed message-authentication code — an HMAC — computed with a secret key unique to the site. This does not hide the data. It makes the data *tamper-evident*: anyone can read it, but no one can alter or forge it without the key, because any change breaks the verification. For a provenance tool this is the more valuable property. The point of embedded ownership information is usually to be seen; what matters is that it cannot be quietly rewritten to name someone else.

The third mode encrypts the record with a strong, authenticated cipher, so that the payload becomes an opaque block only the holder of the site key can decode. This is a true secret, appropriate where the owner wants the embedded record to be unreadable to others rather than merely unforgeable.

The three modes are labelled for exactly what they are. Readable is called readable. The signed mode is described as tamper-proof and signed, never as encrypted. Only the genuinely encrypted mode is called encryption. Independent of the mode, the record can optionally be written under a deliberately ambiguous, innocuous-looking namespace, so that the block is less obvious to a casual metadata-stripper and less likely to be singled out for removal. This is a modest, honest form of concealment — it makes the information harder to notice and strip, not impossible.

Independent, verifiable timestamps

Metadata and watermarks answer the question of whether an image can identify itself. A trusted timestamp answers a different question: whether a particular file, or a cryptographic representation of it, demonstrably existed by a particular time.

The implementation added optional trusted timestamping using the established RFC 3161 protocol. A cryptographic hash of the ownership claim is sent to an independent Time Stamping Authority, which returns a signed token binding that hash to a trustworthy time. The default authority is one of the certificate authorities whose timestamps are already widely relied upon for software code-signing, precisely because such timestamps are routinely accepted as evidence of when something existed; the authority is configurable so that another compliant or higher-assurance provider can be substituted without changing the system.

Two design decisions here follow directly from the honesty principle. First, the feature never fabricates a verified result. If the authority cannot be reached, the system records an ordinary local timestamp and says so plainly, rather than presenting an unverified time as though it had been independently certified. The distinction between a self-asserted date from a local clock and an independently certified timestamp is precisely the distinction that gives the latter its evidential weight, and blurring it would destroy that weight. Second, only a hash is ever transmitted to the external authority — never the image and never the readable ownership data — so timestamping does not expose the protected material.

It is equally important to state what a timestamp does not do, because the earlier technical section already established the point and the implementation does nothing to escape it. A timestamp does not prove authorship. It proves that a particular data existed by a particular time. Its role is to strengthen a chain of evidence, not to complete one.

Verification without the plugin

A recurring test was applied to every evidential feature: if the claim can only be checked by this software, it is not evidence. A protection that says “trust my program” proves nothing to a sceptical third party, because the obvious response is that the program could be mistaken or dishonest.

The implementation was therefore built so that its cryptographic claims are reproducible with standard, publicly documented methods and no dependence on the plugin. The signed metadata uses a standard keyed hash; given the site’s key, any competent examiner can recompute it with ordinary tools and confirm the data is unaltered. The encrypted metadata uses a standard authenticated cipher; given the key, it can be decrypted and its integrity confirmed with widely available software. The trusted timestamp token is a self-contained signed object that carries its own certificate chain and can be verified independently against the issuing authority.

This has a clear consequence for how the system must be operated and described, and the implementation treats it as unfinished work rather than a solved problem. For any of this to serve as evidence in a dispute, the site’s secret key must be capable of being exported and disclosed to a neutral examiner, and the exact

format of what is signed or encrypted must be documented so that an examiner can reproduce the check without the plugin. Until those supporting pieces — a key-export facility and a published verification method — are complete, the honest description of the cryptographic layer is “tamper-evident, cryptographically signed, and independently verifiable with your key,” and deliberately not “court-admissible” or “legally proven.” The stronger language has to be earned by the supporting apparatus, not asserted by the marketing.

A consistent record across the tools

Because these layers share the same underlying facts about an image — its identity, its owner, the copyright statement, the registration details where they exist — the implementation stores that information once, as a single record, and lets each tool read from and contribute to it. Setting a default once means it need not be retyped for every image; the same record feeds the visible watermark’s copyright line, the forensic payload, the embedded metadata and, in due course, the certificate. This is not merely a convenience. It ensures that the different protection layers make the *same* claim about an image, rather than drifting into slightly different versions of the copyright line or creator name across the file. Consistency is itself part of credibility.

Where this leaves the model

The proof of concept proved that acquisition could be made harder. The implementation described here added the identification and evidence stages that the earlier account could only anticipate: visible deterrence, invisible identification, embedded and cryptographically protected ownership information, and independently verifiable timestamps, all resting on the same withheld-master foundation and all recorded against a single consistent provenance record.

None of these stages is complete on its own, and the implementation makes no attempt to pretend otherwise. Prevention is friction, not a wall. Identification survives common redistribution but not every transformation. Evidence strengthens a claim without settling it. What has changed since the earlier account is only that these are no longer descriptions of what might be built. They describe what the system does, together with an explicit statement of what it does not do — which, for a tool of this kind, is the more important half of the description.

The search-visibility problem, discovered late

One consequence of the design only became fully apparent near the end of production, and it is worth recording honestly because it forced a genuine tension in the model.

The central promise of Image Guard is that the public never receives the real image — only an innocuous placeholder, with the master withheld and the displayed version substituted by script for human visitors. That promise is what makes the protection meaningful. But a search engine is not a human visitor. Google’s image crawler does not run the substitution script, and it builds its picture of an image from the file it can actually fetch. When the only fetchable file is a grey placeholder, the protected image effectively disappears from image search. For an artist who protects their work *in order to be seen*, silently removing that work from Google Images is not a neutral side effect — it is the plugin quietly undoing part of the reason the artist published at all.

The uncomfortable truth is that the two goals are in direct conflict. Anything a crawler can fetch to build a thumbnail, a scraper can also fetch. *Indexable as a thumbnail means fetchable*. There is no arrangement in which a real, recognisable image appears in Google Images while remaining genuinely unfetchable. Metadata alone — alt text, an image sitemap, structured data — can catalogue an image without exposing a single real pixel, but it cannot make the *visible search result* anything other than the grey box that was crawled.

The resolution I settled on does not pretend the conflict away. It exposes it, per image, as an explicit choice the owner makes:

- **Maximum protection** — the default, and the original promise unchanged. The only public file is the grey placeholder; nothing real is ever crawlable; the work will not appear in image search.
- **Search preview** — an opt-in in which the public file becomes a small, permanently and heavily watermarked copy of the real image, so image search can show a recognisable but low-value comp. The high-resolution master still never leaves the private store.
- **Abstract preview** — a further-reduced, pixelated, watermarked copy: enough to be found, too degraded to reuse.

The watermark on the preview levels is deliberately not the gentle corner mark an owner might choose for their live display. It is a dense, tiled mesh carrying the plugin's own mark across the whole frame, chosen after testing that removal attempts — upscaling, generative fill — destroyed the subject rather than yielding a clean copy. The preview is designed to be low-value on purpose, because its whole justification is that exposing it costs the owner little.

This is the point at which Image Guard visibly bends its own central promise, and it should be stated plainly: in the preview modes, a reduced and watermarked version of the real image *is* made public, where the default mode publishes nothing real at all. That is a real departure, made only because search visibility is, for many creators, a genuine requirement rather than a luxury. The honest framing is therefore not "the original is never shown" but "the original master is never public, and the owner decides, per image, whether a degraded watermarked preview may be — trading a measure of exposure for discoverability, with the trade made visible rather than hidden."

The guidance that follows from this is simple, and the plugin states it at the point of choice: **if an image does not need to be found through image search, it should stay on the grey placeholder**. The preview modes exist for the images whose discoverability matters enough to justify publishing a deliberately weakened copy — and for no others. The default protects; the exception is a decision, not an accident.

Conclusion: Returning Control to the Creator

The development of RobArt Image Guard began with a technical question, but the reason for asking that question was personal.

My career as an artist began before the internet became an ordinary part of creative practice. I experienced the transition from physical and comparatively controlled forms of reproduction to an environment in which a digital image could be copied, redistributed, altered and commercialised across the world almost instantaneously.

For artists, the internet brought extraordinary opportunity. It allowed work to reach audiences that would previously have been impossible to access. It created new marketplaces, communities, portfolios and methods of communication.

It also created a contradiction that has followed me throughout much of my professional life.

The more successfully an artist shares an image, the more opportunities exist for that image to escape the artist's control.

My own experience ranged from scraper sites and unauthorised merchandise to works being altered, separated from attribution and used in contexts completely contrary to the intentions behind their creation. Attempts to respond through takedown notices, reverse-image searching and copyright enforcement frequently became a repetitive and exhausting exercise. The law could establish that rights existed, but practical enforcement remained constrained by jurisdiction, cost, procedure, anonymity and the simple speed with which another copy could appear.

Eventually I described much of my previously distributed artwork as being **in the wind**.

Once unrestricted copies have circulated widely enough, there is no realistic mechanism by which an artist can recall every version. At that point protection becomes retrospective: find the infringement, identify the infringer, request removal, and repeat the process when the next copy appears.

RobArt Image Guard emerged from questioning that model.

Instead of concentrating entirely upon what could be done after an image had escaped, I began asking whether the valuable image needed to be placed into the ordinary public delivery system in the first place.

The earliest answer was remarkably simple:

| public placeholder, private original.

The implementation that followed developed that idea considerably further. The protected master remains outside the normal public media structure. What the visitor receives is controlled independently, and the public-facing WordPress environment continues to expose the low-value placeholder rather than the protected master.

During the earliest stages of development, I assumed that this could only go so far.

Like most people familiar with browser development tools, I expected that a technically knowledgeable user would eventually inspect the page, locate the delivered image resource and simply open it independently or save it.

That assumption influenced the early description of Image Guard. I did not want to make an unrealistic claim that a browser-displayed image could somehow become completely inaccessible.

Subsequent development produced a result that went further than I expected.

The delivered rendering was changed to use controlled, signed requests and temporary browser-memory objects rather than stable public image URLs. The master itself never enters the public delivery path, and the rendering made available to the visitor is limited to what is required for display.

Most importantly, practical testing of the routes that I had expected to expose the image did not produce the anticipated result.

Right-click saving returned the placeholder.

The public page source exposed the placeholder.

Normal inspection of the public files exposed the placeholder.

Examining the browser's developer tools and attempting to follow the displayed image through the expected source and resource routes did not produce a stable image file that could simply be opened in another tab and saved.

That does not justify describing Image Guard as one hundred percent secure.

Such a claim would contradict one of the principles maintained throughout the project: **the software should never claim more than can actually be demonstrated.**

The pixels displayed on a monitor can still be photographed, screen-captured or reconstructed. New browser behaviours, new attacks or techniques not yet considered may eventually reveal weaknesses. Security should always be treated as something that is tested and improved rather than declared complete.

What can presently be said is narrower, but substantial.

Across the browser acquisition routes tested during development, Image Guard has not exposed a reusable protected-image file in the manner originally expected. The original master remains private, and even the image presented for viewing is not simply sitting behind the page as an ordinary downloadable asset.

For me, that result is significant precisely because I began the project expecting that there would inevitably be an obvious way around it.

For years, my experience of digital protection had reinforced that expectation.

There was always another copy.

Another scraper.

Another marketplace.

Another search result.

Another infringement.

The assumption that the determined thief ultimately wins had become almost instinctive.

Image Guard does not prove that this assumption has been permanently defeated. What it does demonstrate is that the balance can be moved considerably further towards the creator than I originally expected.

The delivery mechanism is only one part of that shift.

The wider system has developed around three complementary ideas already identified in this paper:

Prevention — reduce or prevent straightforward acquisition of the valuable image.

Identification — allow a distributed image to retain or recover a connection with its creator through watermarking and embedded provenance information.

Evidence — preserve hashes, ownership records and independently verifiable timestamp information that can strengthen a later claim.

No one of these is sufficient by itself.

A visible watermark can be removed.

Metadata can be stripped.

An invisible watermark may eventually be damaged.

A timestamp proves existence at a time rather than authorship.

A protected delivery system cannot prevent somebody photographing their screen.

But protection does not have to depend upon one perfect mechanism.

A realistic system can instead make theft difficult, identification persistent and provenance increasingly difficult to dispute.

That is the approach Image Guard has taken.

There is also a broader reason for developing it.

Creative practice is already too often treated as though it is somehow less legitimate than other forms of work. Digital artists, illustrators and photographers are regularly expected to publish material publicly in order to promote themselves, while simultaneously accepting that the same act of publication may expose the product of their work to unauthorised reproduction.

That has consequences beyond hurt feelings.

Artwork represents labour.

For many creators it represents income.

When unauthorised copies are converted into products, supplied through competing sellers or distributed freely in place of licensed versions, the creator is not simply losing control over a file. The economic structure that allows that creator to continue producing work is being undermined.

Over time, the fear of that happening can also alter the creative process itself.

My own response was eventually to watermark heavily, crop artwork, withdraw work from sales platforms and become increasingly reluctant to publish complete images. The irony is obvious: a system intended to allow creative work to reach the world can ultimately persuade the creator not to show it.

That is the problem I most want Image Guard to help change.

The purpose is not merely to stop somebody pressing **Save Image As**.

It is to allow an artist to publish because they want the work to be seen, rather than spending the entire process thinking about how quickly it might be stolen.

Creative Licence, SylentID and the other experiments that preceded Image Guard were all attempts, in different ways, to address the same problem. Image Guard is the most complete expression of that effort so far.

I do not regard it as the final answer.

A more experienced security researcher or developer may find methods of strengthening what has been started here. New technologies will undoubtedly create new methods of attack as well as new methods of protection. Image Guard should therefore be regarded as a contribution and a stepping stone rather than an endpoint.

That is sufficient.

If it helps move digital-art publishing towards a position in which creators can share their work more confidently, then it has served its purpose.

Artists should not be made to feel that their work is **in vain**, nor should reaching a wider audience necessarily mean relinquishing practical control of the property from which their livelihood may depend.

My experience of online image theft is the reason this project exists.

Its objective is ultimately very simple:

to put as much control as technology reasonably can back into the hands of the creative people to whom that control belongs.