

**WHAT WE DON'T SEE**

Christiana Caro

## WHAT WE SEE

What we see is resting on a horizon

Out to sea

Diana's Temple is crumbling around us, pillars of  
stone made strong with aging mussels

A crooked pyramid at 38 degrees rests unsteadily,  
offset by that straight edge, that constant visual  
correction

What we see are three mounds laid out on a line

Rocks rocks rocks

As crystals, gems, fossils, islands, memorial ruins  
to pasts not yet distant enough

The days go by, that's what I wanted...Each tick  
closer to being further away

What we see is a monster lit by a flame

Our mother's hand, a comfort

Fire-set sheep run through the woods, leaving virgin  
land in their wake

Scorched Hephaesti forging mercurial expansion from  
the burn of Saturn's darkness

This year's scirocco blows in oven-hot, with a  
curly-tailed warning sauntering through the square  
just before midnight

We shoot cafe whiskey with an Irish poet and toast the  
possibility of chaos, of leaving, of not knowing what  
ever comes but what has been and is still now, even  
now

The burning wind in our hair picks up, threatening  
that it may only get worse later, we just don't know

And we laugh

What we see are hills splayed in all directions, mist  
rolling in, quickly moving cool and wet

The day's bright sun heating red rocks that hold its  
warmth hours after

Even into the empty crypt shaped to fit filled concrete

Even into the night again

Where we lay down roadside on beds of stinging greens  
and goat shit

Looking

Waiting

For stars, while we compose our three true desires

We get 2/3.

What we see are men gathered at the square, their boys  
practicing along the edges

Looking

Waiting

They sit on the church steps, hovering around the  
tombs or on top

Where holy bodies were hung and drained and rotted and  
piled

Before being sealed under skull-pressed marble, before  
becoming the flat planes upon which weekly worshippers  
greet their altar

What we see is a garden tended by the sisters, who say  
they will pray for us from bare rooms with abundant  
views

Ripe blackberries on the vine

Chosen carefully as we walk further into the Tardara  
forest

Or up to the breast-shaped hill

Or down to the valley below

## CONVERSATION

**Kate Palmer Albers:** Let's start with place. Why did you  
choose Sardinia?

**Christiana Caro:** The project behind my recent work is  
based on research I'm doing around the International  
Latitude Observatories, a series of now-defunct land-  
marks built in the early 20th century that once measured  
the earth's wobble (aka it's "precession") on its axis  
from the same point of latitude in six locations around  
the globe: California, Italy, Maryland, Turkmenistan,  
Ohio, and Japan. So far, I have made images at half of  
the sites. The observatory in Sardinia is both an island  
and a local museum – two qualities that appealed to me.  
My two-year contract with Google had just ended, and I  
was excited to go out into the field with the camera eye  
that I had dedicated a significant amount of myself to. I  
was accepted into the Officina Stamperia Del Notaio art-  
ists residency, located in the remote Sicilian mountain  
town Tusa, where I spent about a month in 2018.

**KPA:** So, these observatories all were built along the  
same degree of latitude, 39° 08' north. They watched the  
same stars to understand how the earth was moving. Aside  
from the science, that's a pretty compelling proposi-  
tion. What about the earth's wobble was interesting to  
you?

**CC:** The wobble describes a shift in the axis of Earth's  
rotation. It's a massive geological event, but describ-  
ing it as a "wobble" sounds so minor and inconsequen-  
tial, even silly. I'm interested in that difference,  
between an event and its name. And that we've come to  
depend upon that wobble, and its measurement, so deeply  
because it's the source of modern GPS.

**KPA:** Which we rely on for everything, it sometimes seems  
like. So many aspects of our daily lives – from the  
maps on our phones to the UPS delivery route, to our  
sky chart and moon apps – use GPS technology. We rarely  
think about the system it's connected to, and how recent  
the technology is. How did you learn about the observa-  
tories?

**CC:** I became aware of them while doing research around observatories on islands that I went looking for in Slovakia and Finland about ten years ago. There was just something kind of compelling about finding all of the latitude observatories, connecting those dots around the globe. They bridge pretty directly to my earlier work, like the 10-mile points project, which was all about observation, measurement and the landscape.

**KPA:** For those 10-mile point photographs, you measured the distance from your apartment in Boston, MA, ten miles in each cardinal direction: North, South, East, and West, and also included Northeast, Southeast, Southwest, and Northwest. You know I love this series, starting from the concept that the locations were both so precise and so arbitrary, and, until you got to the site, also so unknown. Once you got there, wherever "there" was, you had to figure out how to make an interesting photograph. It's a great set of aesthetic constraints, the kind that actually allows a lot of freedom. I can see how this set of observatories, all lined up, and looking toward the same stars, resonates.

**CC:** Instead of measured points on a local map this project used landmarks around the globe, but yes I think those impulses are closely related. A lot of my work has been about designating these firm points, but using the route to get there to make the images. People like Richard Long, Hamish Fulton, Ana Mendieta – early conceptual artists creating work about movement or mark making in the landscape – really informed my development as an artist. The landscape is central here. When I got going on the 10-mile points, in 2000-2001, GPS was not yet commonly used, and the technology was fairly rudimentary. I went on these roundabout and weird adventures to get to the points. I would get lost. It was so much about the in-between, not just the endpoint. It was kind of a performance of making images, of moving through the landscape in that specific way.

In terms of the observatories, each one is a designated point. They relate to each other – from huge distances – in such a specific way, connected and reciprocal. Yet, each site has its own distinct identity. The Sardinian architecture, for example, is informed by a North African influence, whereas the one in Ukiah, California is a



GC\_20180729\_104544\_1533064044.jpg



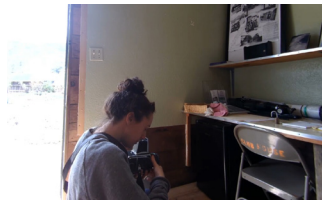
tmp\_100\_n7nYSv.jpg



GC\_20180725\_163443\_1532543476.jpg



IMG\_5666.jpg



tmp\_62\_8gDR03.jpg



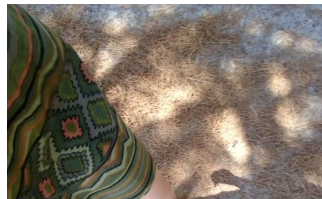
GC\_20180728\_103359\_1533065059.jpg



GC\_20180725\_130936\_1532543051.jpg



tmp\_66\_6JtcpF.jpg



GC\_20180729\_110353\_1533063197.jpg

white clipboard box. Each one reflects a regional architecture. And, of course, getting to each one involves quite a process. So yes, these projects definitely come from a common curiosity of mine, which is to see what will happen within a given framework, particularly one that takes me on some kind of wandering orbit.

There's something else to add here, which is that I was also thinking about the precession of the equinox, this slightly shifting movement of the atmosphere that results from this wobbling Earth on its axis.

**KPA:** I don't know about this.

**CC:** It relates to observation, and place, and the illusion of a fixed reality. The precession of the equinox refers to the phenomena of the rotation of the upper atmosphere, a cycle that spans a period of about 20,000 years, over which time the constellations appear to slowly rotate around the earth. Over time it affects the patterns we see, what is observable from earth. There are ancient constellations, such as the Argo Navis, that were recorded, but that are either gone or intermittent-

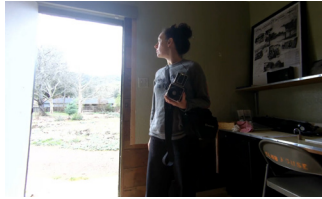
ly fading from view. They've taken turns appearing on the horizon.

**KPA:** The scale is almost incomprehensible. Let's turn to your time working for Google, and how that ties in. You were hired to help train an A.I. camera, called Clips, that takes pictures when it "sees" what it has learned is a good picture.

**CC:** Right. In 2015, Google hired Andy Rogers, a photojournalist; Christopher Woon-Chen, a filmmaker; and me, a fine art photographer, to teach an autonomous, artificially intelligent camera about photography. We were hired to fill it with directives – this is good, that is not. Yes, no, yes, no. And it was not only about what made a good photograph, aesthetically, but also about meaning. What I found from two years of working on the project was that, although we could train the camera about composition, the rule of thirds, to ignore "obfuscation" of the camera lens, etc. in the end, it was an incredibly complex task to imbue the camera with a sense of when the rules should "break." This is something I teach my photography students after teaching them the basics of



tmp\_93\_UnifDf-3.jpg



tmp\_63\_0i7bTn.jpg



GC\_20180726\_161444\_1532627419.jpg



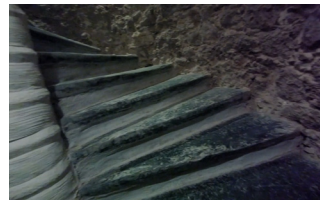
GC\_20180729\_105838\_1533063882.jpg



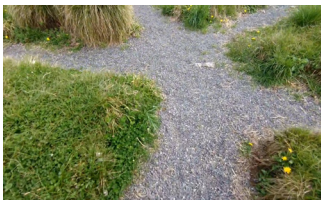
GC\_20180729\_110309\_1533063217.jpg



GC\_20180729\_104413\_1533064113.jpg



GC\_20180728\_110427\_1533064802.jpg



tmp\_93\_UnifDf-2.jpg



GC\_20180729\_113023\_1533065265-2.jpg



GC\_20180725\_160831\_1532543030.jpg



tmp\_87\_5FH6vR.jpg

composition. But it's a much more difficult concept to explain to a machine.

Beyond aesthetics - the what - some of the largest challenges we ran into were around deconstructing "meaning" into entities that a camera could learn to recognize and see - the when and why. Through research, we learned that meaningful moments were often comprised of smiles, familiar people, and interactions between those people. So the camera was programmed to look for upturned lips, bodies touching or in movement, and a capability to discern which people in the frame were important to the user. However, what the camera could never learn is that meaning is also constructed from what you cannot see: the context of a gathering, the why behind a smile, the significance of two people hugging, how that moment felt. A lot of my personal work comes from these kinds of suggestive rather than descriptive places. This occurs even within the context of rigid project restraints based on parameters such as specific monuments, land forms, or measurements. The ephemeral always surfaces.

**KPA:** You're talking about describing, or breaking down,



GC\_20180729\_104553\_1533064093.jpg



GC\_20180729\_111919\_1533063537.jpg



tmp\_85\_sjhfV1.jpg



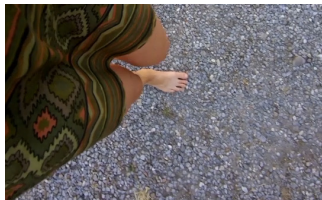
GC\_20180729\_104526\_1533064024.jpg



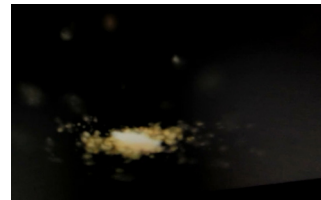
tmp\_70\_z95d5W-2.jpg



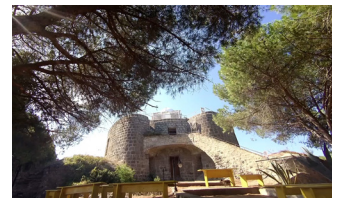
GC\_20180729\_105702\_1533064339.jpg



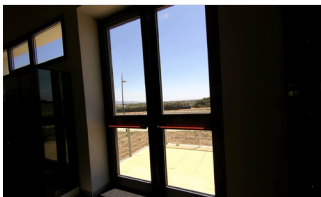
GC\_20180729\_110256\_1533065552.jpg



GC\_20180726\_111027\_1532627344.jpg



GC\_20180728\_102002\_1533064742.jpg



GC\_20180725\_132712\_1532543322.jpg



tmp\_59\_fE6TzZ.jpg



GC\_20180729\_104705\_1533064189.jpg

a process of understanding visual language, and the language of gestures, that is deeply intuitive to humans. What we notice, how we see, and how we make meaning. You taught the camera to see, essentially. You summarized the strangeness of this process really well in the title of a talk I heard you give "Teaching Machines to See, or, That Time I Uploaded My Brain to an A.I." It strikes me as incredible, on the level of individual experience. What is the bigger picture here?

**CC:** Fundamentally, to automate the detection of a desirable image is to consider the nature of observation itself. This automation of the gaze, built on expert human instruction, is a kind of collaborative perception. But there is something of the void in human visual experience that maybe cannot be reproduced by a machine, or is not the instinctual way of approaching one.

Here's an example: While working on Clips, I started noticing yachts floating by on the canal outside the window of my desk. My colleagues and I created a frame out of scotch tape and cardboard, attached that "frame" to the window, and titled the experiment "Yacht TV". It was a

nod to Norway's slow TV movement, and it was meant to be absurd in a way, originally. The beauty of this was that people across the office started to ask what the experiment was. Eventually, I realized, people would come by my desk specifically to enjoy the experience of noticing, of watching boats pass by within the frame. It was a reprieve from their screens.

To push the project a little further, I began shooting videos on my phone when boats passed by. I then decided to make it available more publically, by publishing "Yacht TV" moments to Instagram. Then, an engineer and I trained an AI to identify the boats and post automatically to a Twitter feed. He taught the AI the elements and styles of boats to look for and identify, and to "wake up" when it detected a moving object. But it was like what I learned working on Clips: The meaningful part happened in what wasn't captured or posted to a feed, i.e., how important the mundane persistence of flat water actually was to the overall experience of finally seeing a boat float through the frame. I started to think about seeing for seeing's sake versus seeing as a means for capture.

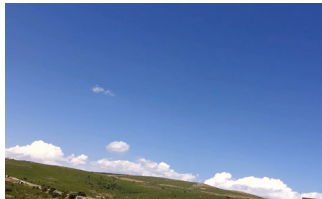
This contrast is something I've continued to think about in my own work. What do we see? What is empty, but still critically important? Where, in these experiences, is the meaning?

**KPA:** These questions seem simple, in some way, but also connect to the history of the relationship between human beings and cameras, particularly in a creative realm. It used to be that, for the most part, critics questioned whether photographs could ever be art, because the camera is, by nature (so it seemed), a mechanical eye - not a human one. But this work, training the camera to see like a person. does it make the camera more human? Or does it underscore the gulf between human vision and mechanical vision? Hasn't photography always involved this kind of collaborative perception, with the human understanding how any particular camera sees? Does this just shift the timing of that collaborative perception?

**CC:** You're collaborating, kind of. The spirit of human collaboration is two entities with individual knowledge coming together to create a third thing. But in this case, the camera does not bring personal experience or



tmp\_21\_XHjoEs.jpg



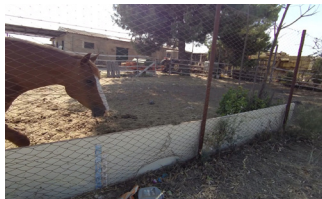
GC\_20180725\_160012\_1532543370.jpg



GC\_20180729\_113023\_1533065265.jpg



GC\_20180725\_161940\_1532552025.jpg



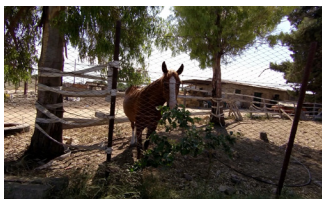
GC\_20180726\_161623\_1532627331-2.jpg



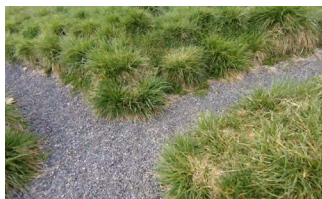
tmp\_84\_ggRqTz.jpg



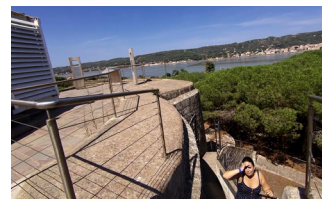
tmp\_47\_3gAMMx.jpg



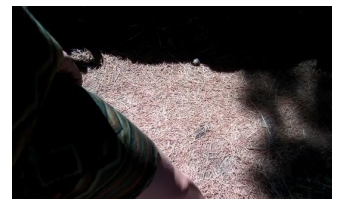
GC\_20180726\_161444\_1532627419-2.jpg



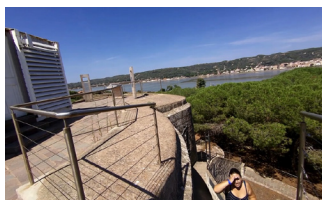
tmp\_90\_b1CVfg.jpg



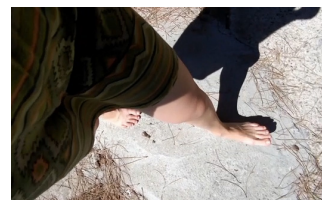
GC\_20180728\_113614\_1533064631-2.jpg



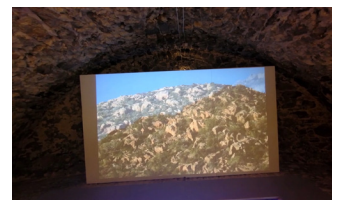
GC\_20180729\_110126\_1533063714.jpg



GC\_20180728\_113614\_1533064631.jpg



GC\_20180729\_110329\_1533063178.jpg



GC\_20180728\_103519\_1533064677-3.jpg



GC\_20180729\_110335\_1533063182.jpg



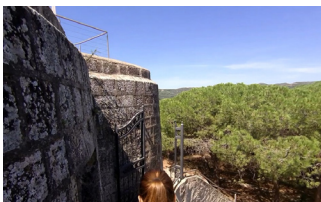
GC\_20180729\_110231\_1533065505.jpg

bias, it's only reacting based on what directives it has been given. Its collaborative ability is limited by the level of technology required. One version of the A.I. camera – which hasn't come to pass, partially because of the complexity of the engineering involved – is that the camera would actually learn, would respond to its user, would potentially become a kind of collaborator. The camera would pay attention to: What do you look at? What do you like? What do you want to see? And modify its actions accordingly. That's the ideal, and it's a cool thing to imagine, but it's not quite happening. Yet.

You know, it's interesting that in my interview with Google, before I was hired, I told them about my approach as an artist. This included my interests in mapping, narrative, and landscape, how I always shoot film (that I sometimes don't see the outcome of for months), that I'm an analog dinosaur, that everything is instinct. They wanted the spirit of that in this product. They were interested in photo history and, especially, in Henri Cartier-Bresson's idea about the decisive moment. That was key. And what began as a month-long assignment to create a few "reference videos" using

machine perception to extract meaning from data turned into a 2-year deep dive into the nature of seeing and image capture. The wild experience of artists and engineers being thrown together to create a language around transcribing aesthetics into quantifiable code came along with that, and ended up sparking incredibly self-reflective moments around why we (as artists/visual thinkers) responded the way we did to photographs. Or chose one frame over another. Nothing was taken for granted.

The project was so much about distilling knowledge into information, into what moment to capture: to know where to position yourself, to know when to capture the image, and how to augment human memory and storytelling. And then, to put it all on a mobile device. The roots were genuine. But at the end of the day, it's very hard to train a personal aesthetic, and especially hard to sequence a narrative, even with a team of some of the best machine learning engineers in the world. That's actually really hard to do as a human, let alone translate into code. It seemed easy enough at the outset, but when you get into the nitty gritty, why this and not that? What are the transitions? It's so complex. So it got simpli-



GC\_20180728\_113633\_1533064646.jpg



tmp\_70\_z95d5w.jpg



GC\_20180729\_105804\_1533064335.jpg



tmp\_74\_rPAGwI.jpg



GC\_20180725\_125620\_1532543150-2.jpg



GC\_20180728\_103453\_1533064987.jpg



GC\_20180726\_161623\_1532627331.jpg



tmp\_88\_QVye9.jpg



GC\_20180726\_161507\_1532627470.jpg



GC\_20180725\_155852\_1532543361.jpg



fied down to: "Smile." "Movement." "Person." "Touch."

**KPA:** Last year you wrote: "As humans, we instinctively engage with the natural world and are destined to continue extending our bodies through technology... As we become interconnected with technology, we are not evolving and adapting, or being replaced, dominated or destroyed. We are being seen." Having finished the Clips project and now made photographs alongside the camera, looking at subjects also about observation, do you feel the same way?

**CC:** I'm interested in parsing these varied modes of observing (machine eye, human eye, architectural eye) and, the idea of self observation, in the sense that we are revealed through our evolution alongside of our technology as we understand who we are and what it is not. On this trip, I had just completed my work with Clips. I knew the energy, the excitement, and the politics inside the sort of ground zero of technology today: Research and Machine Intelligence at Google. Some people joked with me when I was working on Clips that I was replacing myself, making myself as a photographer obsolete.

But the nature of developing technology is that as it advances, so do we. And we then create and program new technologies. It is a continual feedback loop, one will not replace the other. We are interdependent. I also knew, or still know, that this technology is going to travel the same path of the observatories. It will become defunct, even absurd to reflect back on someday. It's hard for us to imagine our own present, which seems so technologically advanced, becoming outmoded in this way, seeming absurd to people in the not too distant future. But of course it will happen. And then we will laugh at these earlier versions of ourselves, the ruins. And all the while, our soft in the middle wobbling megolith of an Earth will keep on turning, too.

**KPA:** You know I had a less smart version of a camera like this. For a few months in 2012 I wore around something called the Memoto camera - I got it on Kickstarter. It was a little orange square that I attached to my shirt and it automatically took a picture every 30 seconds. There was no A.I., the camera wasn't trained. It just took a huge volume of pictures, indiscriminantly; it was part of the lifelogging trend. I have thou-



tmp\_67\_iQ6NLf.jpg



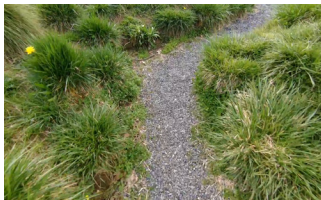
tmp\_98\_4RnhaZ.jpg



IMG\_0474-2.jpg



GC\_20180725\_155426\_1532552005.jpg



tmp\_89\_nmIA1J.jpg



tmp\_48\_NNNtpI.jpg



GC\_20180726\_112152\_1532627496.jpg

sands and thousands of pictures that this camera made, as I wore it around Tucson (where I was living then): to class, to social events, to restaurants, driving around town, whatever. I always wanted to write something about them, but the pictures are so boring. It's a lot of skies and ceilings, the inside of my car, or the edge of a table, or off-kilter frames that maybe include the person I was talking to, but, more often, don't. The main things I learned from this camera were 1) how narrow my view of what I "see" is (almost totally focused on people and what is directly in my line of sight) and 2) how more pictures definitely does not add up to more information. It also, now that you mention the timeline of the new becoming the newly defunct, seems totally dated already. That's a long way of getting around to my question: How do you reconcile the banality of the Clips pictures with the kinds of images you make as an artist?

**CC:** I had this idea about collaboration, but in the end I'm not sure what it did for me, image-wise. The technology is just too rudimentary, and more importantly, it takes time to switch tools after 20 years. My eye has evolved by seeing with my particular tool (a Hasselblad

film camera), which in turn has resulted in my personal style of image-making. The two are somewhat indistinguishable now. If anything, I considered Clips as more of a note taker, maybe a partner in research. As I moved through different locations, I often had Clips passively recording alongside of me. When I went back to look at the images, it felt more like scouting footage, to remind me of where I had been. I prefer to shoot alone, but to some extent I did feel like somebody, or something, was out there with me this time. Mainly, working on this camera made me think about seeing constantly. and, especially, the differences between seeing and capture. Cameras used to be disconnected eyes, and now they're beginning to "understand" what they see, to make decisions, and someday they'll start learning from their mistakes. It is undeniable that A.I. is creating a new language around not only how cameras work, but also influencing the kind of images we make. That being said, I believe that these relationships with our technology are also distinctly personal, changeable, and constantly in flux.

**KPA:** One of the things you said in that talk, which re-



GC\_20180729\_105702\_1533064339-2.jpg



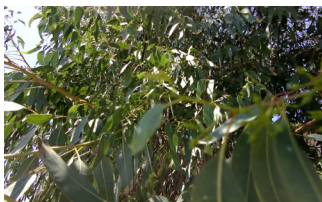
GC\_20180728\_103519\_1533064677-2.jpg



GC\_20180729\_110322\_1533063171.jpg



tmp\_44\_0rDcwG.jpg



GC\_20180725\_125620\_1532543150.jpg



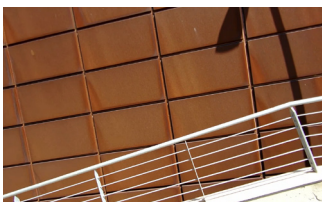
GC\_20180725\_160839\_1532543043.jpg



GC\_20180729\_113939\_1533065141.jpg



tmp\_97\_h2dp80.jpg



GC\_20180726\_163510\_1532627505.jpg

ally resonated with me, was: "I believe AI is a new kind of tool, and as artists we have an opportunity to play with it, and a responsibility to understand it." Do you still feel this way?

CC: Yes, I absolutely do. What are the possibilities of A.I. in art? What might developments in machine intelligence, neural networks, or artificial intelligence mean for human creativity, or collaboration? These are open questions, and they are important ones. At any rate, it is our inevitable future, so we may as well join the party. As artists we have a massive responsibility to understand and respond to the forces that shape our culture. The work we make has the potential to translate these technological paradigm shifters into narratives that will in turn affect the development of the technology itself. This is a moment of opportunity to PLAY with, to incorporate into our practices, and even to emerge in tandem with our A.I. future.

## COLOPHON

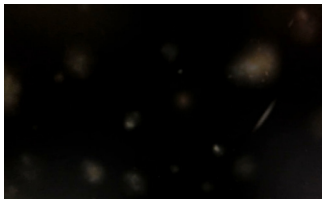
\*\*\*\*\*



GC\_20180728\_103519\_1533064677.jpg



GC\_20180729\_110158\_1533064593.jpg



GC\_20180726\_111017\_1532627614.jpg



tmp\_83\_eaMzuI.jpg



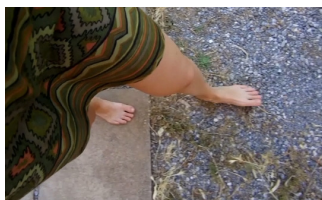
GC\_20180729\_110316\_1533063221.jpg



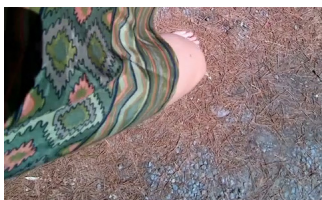
GC\_20180725\_125353\_1532543135.jpg



tmp\_19\_1hK7J5.jpg



GC\_20180729\_110302\_1533065557.jpg



GC\_20180729\_110238\_1533065518.jpg



tmp\_94\_X5PRsc.jpg

This pamphlet was created on the occasion of **WHAT WE DON'T SEE**, featuring work by Christiana Caro, at the Greenleaf Gallery, Whittier College, September 5 to October 11, 2019.

The images were taken by the Google Clips camera, about which Caro has said:

Clips is an intelligent camera that I was hired by Google to teach about seeing. I was instructed to fill it with my own, my expert, aesthetic bias, and what I deemed as meaningful within a stream of recorded moments. This is good, this is not. Yes, no, yes, no.

The images were placed in the pamphlet using an online random number generator which determined how many images would be on each page, and where in the grid of ten each of those images would be placed. The ordering of the images was also computer randomized, ultimately including 84 of the 104 images originally chosen for possible inclusion by Caro.

The accompanying interview between Caro and art historian Kate Palmer Albers is based on a series of phone conversations they had discussing the relationship of Caro's creative work to her experience developing Clips at Google.

The poem **WHAT WE SEE** is by Christiana Caro and published here for the first time.

The pamphlet is set in Noto Sans Mono, a contemporary monospace font like those typically used when writing computer code. It was developed and released under an Open Font License by Google.

