About Us Archives

chives Section

Open Source

Despite popular perceptions that China is a black box, creative new research methods are shining a light in.

BY KATRINA NORTHROP - JANUARY 16, 2022



It's a Friday morning in early January when <u>Adrian Zenz</u> gets up from the desk in the corner of his bedroom to turn up the heat. It's 17 degrees below zero outside, and his quiet neighborhood in Minnesota's Twin Cities has been blanketed in a thin layer of snow.

Dressed in a button-down shirt with rectangular glasses and cupping a warm mug of coffee, Zenz doesn't look like much of a rabble rouser. A German evangelical Christian, his plans for the weekend include taking his "pandemic-puppy," a Shih Tzu named Rosie, for a hike by a nearby river and eating dinner with his family.

But thanks to Google and his desktop computer, the 47-year-old researcher has uncovered information about human rights abuses against Uyghur Muslims in China's northwest region of Xinjiang, which in turn has made him a frequent target of China's ire. Last year, for instance, he was <u>sanctioned</u> by the Chinese government and <u>sued</u> in a Xinjiang court for "fabricating 'forced labor' <u>rumors</u>." On the website for China's Foreign Ministry, prominent dissidents like <u>Liu Xiaobo</u>, the late Nobel Peace Prize winner, and <u>Joshua Wong</u>, the Hong Kong activist, garner around 200 hits reproaching their "anti-China" activities. Zenz gets more than 400 such admonishments.

"Researchers don't become celebrities," says Zenz, his German accent prevalent. "But they [the Chinese government] talk against me almost daily."

Zenz became uniquely high profile starting in 2018. Despite living 6,000 miles away and only visiting Xinjiang once — as a tourist 15 years ago — Zenz's research is fundamental to claims that the Chinese government's actions there amount to a form of



genocide, and his findings have been \underline{cited} by the U.S. government.¹

Adrian Zenz. Credit: <u>Newlines Institute for Strategy and</u> Policy

Mining data from Chinese sites, he has documented

the repression and <u>securitization</u> of the region, the extent of <u>detentions</u> in re-education camps, Uyghur forced labor <u>programs</u>, and most recently, troubling birth control <u>policies</u>. With fluent Chinese and a personality he describes as "dogged," Zenz has also authenticated large leaks of official documents, including the <u>Karakax List</u>, and analyzed Xi Jinping's speeches in the recently published <u>Xinjiang Papers</u> for clues about the central government's policies.

His research, says James Millward, a leading authority on Xinjiang at Georgetown University, has been "creative, systematic and convincing" — an impressive feat considering Zenz doesn't have the backing of a prominent institution or access to high ranking sources inside the Xinjiang government.

"I got really addicted to finding evidence. For me, the thrill is to find something new, and I search until I find it," Zenz says of his approach. "People don't realize what's online."



出境未归人员亲属送培学员

A redacted document from the "Karakax List," a leaked document analyzed by Zenz and the *Financial Times* which appears to document the surveillance and imprisonment of hundreds of individuals from the Karakax region in Xinjiang. *Credit: <u>The Journal of Political Risk</u>*

Because of discoveries by Zenz and other researchers like him, that is starting to change. With China's political repression silencing sources, policy initiatives like the <u>data security</u> <u>law</u> stifling cross-border information flows and pandemic-related travel restrictions, which make on-the-ground information gathering nearly impossible, many researchers are turning to open source methods to keep tabs on what is happening inside of China.

Recent examples of this type of work run the gamut: researchers have used government procurement documents to study the Chinese military's <u>use</u> of AI, satellite imagery to <u>identify</u> Chinese nuclear silos, machine learning to <u>measure</u> coal emissions from Chinese factories, and patent filings to <u>learn</u> about China's development of a surveillance state. Journalists also use these methods, most notably *BuzzFeed News*' Megha Rajagopalan, who shared the international reporting <u>Pulitzer Prize</u> last year for her <u>investigation</u> into Uyghur

detention camps in Xinjiang using satellite imagery.

Despite popular perceptions that China is a "black box," these successes show just how big the window still is for research.

"China is a very digitized society," says Xiao Qiang, the founder of China Digital Times and a research scientist at the University of California, Berkeley, who is using procurement records to map out Chinese surveillance. Government portals, company websites, and social media platforms are "a very rich resource if you know what data you are looking for and ask the right questions," he says.

For <u>Christian Göbel</u>, a professor at the University of Vienna, open source research has been something of a revelation. Göbel used to rely primarily on interview-based research, but after Xi Jinping came to power, his interviewees started clamming up and no longer offered substantial insights. Since then, Göbel has shifted focus to publicly available data. Recently, he tapped Chinese government complaint portals to study the government response to popular sentiment.

"It is remarkable that this is possible in a regime that is so closed," he says. "Just pick any local government website, and you can find really interesting information there."

It's a realization that even the U.S. government is starting to get behind. For instance, the Defense Intelligence Agency (DIA), which specializes in military intelligence, is planning to lead an effort aimed at using open source data to combat threats from countries like China and Russia, according to a <u>report</u> in *The Wall Street Journal*. <u>Robert Ashley Jr.</u>, a former director of the DIA, says the move shows the intelligence community's increasing recognition of the value of open source information.

6 We are moving past the cultural stigma that unless it comes from a classified source, it isn't worth it... There is a lot we [the U.S. government] are leveraging with open source, but there is more we can do. The policy is a bit lagging.

- Robert Ashley Jr., the former director of the DIA

"We are moving past the cultural stigma that unless it comes from a classified source, it isn't worth it," Ashley says. "There is a lot we [the U.S. government] are leveraging with open source, but there is more we can do. The policy is a bit lagging."

Emily Harding, a senior fellow at Center for Strategic and International Studies in Washington and former deputy staff director at the Senate Select Committee on Intelligence, also says the U.S. government is behind the curve. "There are lots of private organizations out there competing with the intelligence community on this," she says.

There are also private individuals leading the effort to tap open source data. Though Zenz is trained as an academic, with a PhD from Cambridge University where he studied Tibetan education, he began researching Xinjiang as a side project, while doing IT work and teaching a research methods course at the European School of Culture and Theology in Korntal, Germany. It wasn't until 2019 that he became a senior fellow at the Victims of Communism Memorial Foundation, a Washington, D.C.-based non-profit, which enabled him to do research full-time.

With a hard drive meticulously organized into folders on specific topics — just in case a web page is taken offline by the authorities or becomes inaccessible to people outside China — Zenz represents the future of research on China. But as business people, politicians and

academics come to depend more and more on this type of work, observers also warn that there are risks.

"People have been very creative in exploiting the Chinese internet as much as possible," says <u>Andrew Fischer</u>, a professor at Erasmus University Rotterdam and an expert on development in western China. "[But] we have to be very cautious with it and take it with a huge grain of salt."

"DIGITAL EXHAUST"



Satellite imagery from Google Earth allowed Robert Barnett, a Tibet scholar at SOAS University of London, to identify Chinese construction activity along the Tibet-Bhutan border. Credit: Robert Barnett/Google Earth

C hina's lack of transparency, observers note, is not a particularly new phenomenon, nor is the need for China researchers to think creatively in order to glean insights about it. In the 1960s, for instance, CIA agents would stalk the Hong Kong fish markets in order to collect the dirty <u>newspapers</u> that wrapped imported fish from the mainland. And researchers studying Tibet have long trumpeted the importance of alternative research methods — both because of Tibet's relative inaccessibility and in order to protect the people on the ground.

But, over the past three decades, as the broader China studies community started to enjoy relatively good access to the country, <u>Robert Barnett</u>, a Tibet scholar at SOAS University of London, says many academics were dismissive of open source methods.

"We weren't ignored by China studies people, we were humiliated by them," Barnett says of his fellow Tibet scholars. Now that other parts of China have become more restrictive, he says that has changed. "We are watching the China field get closer to the Tibet field," he says.

Last year, Barnett published a blockbuster report in *Foreign Policy* showing that China is expanding its territory across the Tibet-Bhutan border by building villages on Bhutanese land. His team stumbled upon the discovery, Barnett says, through a series of Chinese press reports describing how an important official had made a long journey to a remote village. Barnett says that it caught the team's eye because they had never heard of the village, and because the reports mentioned that they needed to cross a specific mountain pass to reach the village. A historian on Barnett's team immediately recognized the pass as being situated along the Chinese-Bhutan border, which set the team off on an investigation, including gathering satellite imagery as well as Chinese and Bhutanese government and media reports.

"No one single piece of evidence will be a smoking gun, but an article can give you specifics about a policy and you can piece different things together," Barnett says of the approach. "Then, you go to satellite photos, for example, to confirm it."

Zenz describes a similar process of triangulation. His efforts often begin with finding the exact terminology for a research subject — which is harder than one might expect due to government jargon — and then plugging those keywords into an online search engine and

meticulously collecting and archiving all of the Chinese government documents that pop up. In 2018, for example, this process led to a <u>trail</u> of 73 procurement bids related to Xinjiang reeducation facilities, which were valued at about \$100 million.

Once he uncovers an important piece of evidence like that, Zenz sets about corroborating it elsewhere — such as with government documents, state media reports or witness statements — to analyze its significance.

"The key to my method is surrounding the topic from different angles and on different levels," he says. "I try to have some evidence that this is a wider policy or practice, then some more detail and some evidence of local implementation, then maybe specific individual cases from propaganda accounts."

This type of cross-government corroboration is possible, observers say, because of the sheer size of the Chinese government, which leaves behind a digital paper trail.

"The Chinese government is a sprawling organization," says <u>Charity Wright</u>, a former National Security Agency employee who is now a cyber threat intelligence analyst at Recorded Future, a cybersecurity firm near Boston. "They like to maintain the appearance that they have pretty good control over their different branches and their provincial governments. But the truth is, like any large organization, it struggles with operational security."

In interviews with *The Wire*, many researchers expressed concern that open source data collection may prove more challenging in the future, since the Chinese government has begun making efforts to delete data, scrub official webpages, and limit access to press reports.

Ryan Fedasiuk, a research analyst at Georgetown's Center for Security and Emerging Technology (CSET), says that he goes to great lengths to make sure that the open sources he uses do not get burned. In a recent report, *Harnessed Lightning*, which reviewed 66,000 government tenders to understand the Chinese military's use of AI, Fedasiuk's team built a web scraper that would only operate during Chinese business hours and only from Chinese IP addresses, to make it more difficult to identify.

"There are reasons why the Chinese are cracking down on the publicly available information," Fedasiuk says. "They publish information that you wouldn't think they would, and the Chinese are starting to worry about that. So we want to take measures to protect those sources."



Ryan Fedasiuk, a research analyst at Georgetown's Center for Security and Emerging Technology. *Credit: CSET*

Like many of the researchers *The Wire* reached out to, Fedasiuk would not identify the exact portal or site he used, in order to protect future research. But despite his meticulousness, the online source he used to get the procurement records was taken down a few months before he published the report. "We think it was because some people were careless about using it," he says.

Even with the Chinese Communist Party's crackdown on publicly available information, however, <u>Robert Potter</u>, who co-founded Internet 2.0, a cybersecurity firm, says there will always be some "digital exhaust."

"China's internet hemorrhages data on every given day," he says. "For example, camera systems leak out all over the internet and they get indexed by search engines."

To assess how China is adopting AI, this report analyzes a sample of purchasing information published directly by the Chinese military in 2020. In addition to capabilities identified in procurement contracts, the authors draw on theoretical writings and research papers by PLA officers and defense industry engineers to assess how the PLA may use the AI systems it is purchasing, and how these systems fit into its concepts of operations.

CSET's corpus of PLA procurement tenders spans 66,207 records published between March 30 and December 1, 2020. These tender notices run the gamut from technology requirements and requests for proposals (RFPs) to announcements of equipment or software contracts that were awarded to Chinese companies. Different types of procurement information reflect different steps in the PLA's technology acquisition process:

Excerpt from *Harnessed Lightning*, a report by the Center for Security and Emerging Technology which examined tens of thousands of public People's Liberation Army procurement tenders. *Credit: <u>CSET</u>*

The Chinese government may also decide to leave some valuable information public as a calculated trade-off, notes Charity Wright at Recorded Future.

"Oftentimes, as a security practitioner, I can advise a company to secure a database or shut down a social media account, but they have to decide whether they want to reduce functionality in favor of security," she says. "Running such a large government, China may choose functionality over security on some things."

Government tenders, for example, which were invaluable to Zenz's early work, will likely continue to be public since they are designed to ensure that the government gets the best deal from government contractors while reducing the potential for corruption.²

"They [the central government] have taken open government pretty seriously because they want to keep local governments in check," says University of Vienna's Göbel. "If the tenders were closed, then rigging would increase again. The government has become more repressive, but it has also become more participatory in that very narrow sense."

More broadly, the Chinese government also wants to trumpet its own narrative internationally, which means leaving bread crumbs of information for researchers to collect and analyze.

"They have an urge to stop us from knowing, but they also have an urge to publicize domestically and internationally how successful they are," Barnett says. "They want to boast and they want to conceal — researchers are parasites who live in that gap."

"A NEW DISCIPLINE"

G iven the growing number of "parasites," however, observers are cautioning against the pitfalls of open source analysis. Researchers *The Wire* talked to, for instance, pointed to the aftermath of the Boston Marathon bombing in 2013, when a team of novice internet sleuths <u>misidentified</u> the perpetrator using publically available images and video.

"Some open source intelligence is incredibly solid," says <u>Peter Singer</u>, a senior fellow at New America, the Washington-based think tank, and an expert on the future of war. "Other times, there might be early stage speculating or unfiltered discussion that might be erroneously labeled as open source intelligence."

China-focused open source work, analysts note, is especially ripe for misinterpretation because of language barriers, nuances in a political system that is very different from Western-style democracies, and growing U.S.-China tensions. "The data analysis part is much harder than obtaining the data itself," says Xiao at China Digital Times.

• Some open source intelligence is incredibly solid... Other times, there might be early stage speculating or unfiltered discussion that might be erroneously labeled as open source intelligence.

Göbel, from the University of Vienna, also warns about the missing data: real life people.

"What I am looking at is stuff happening on the internet — it is not about citizens going about their daily life," he says, referring to his research. "We can learn about what Chinese citizens see when they move around the internet, but even as we light up these parts of the black box, other parts — like actually living in China — are getting darker and darker. These worlds are apart from each other, and we need to know about both of them."

Given the challenges of on-the-ground research right now, however, open source research is increasingly being used to shape the policies of governments outside of China — for instance, justifying sanctions — and many researchers say they need to be committed to getting it right. The increasingly hawkish stance towards China in the West only adds to this pressure.

"In the current climate, over-interpretation and reading something that you want to find such as internal conflict or forced labor — that is the risk," Barnett says. "There are a lot of holes and gaps in what we do."

Even Zenz's work has been criticized. In his recent report on sterilization and birth control policies, for example, some critics say he extrapolated beyond what is established fact and cherry-picked evidence to suit the conclusion. Others, including the Chinese government, have said he is overly influenced by his faith.

Zenz, for his part, says he has made corrections to his reports after errors have been pointed out to him, but he takes issue with the idea that his faith has distorted his research conclusions, as if, he says, it's "some kind of right-wing Christian fundamentalism."

"I feel used by God to work for good, to uncover an atrocity, to work for justice," Zenz says. "For a long time, I have had to do my own thing, figure it out on my own and develop my own method to break this open. This has just been one exhausting pioneering effort."

6 I have talked to governments, but they have admitted that they don't have the manpower to do this and they rely on my work...it's ridiculous that the biggest superpower in the world relies on people like myself for much of their research and insight.

"

— Adrian Zenz, a senior fellow at the Victims of Communism Memorial Foundation

Although Zenz is committed to the effort, he also expresses frustration that so much of this burden falls to individuals like him.

"I have talked to governments, but they have admitted that they don't have the manpower to do this and they rely on my work," Zenz says. "It's ridiculous that the biggest superpower in the world relies on people like myself for much of their research and insight."

On some subjects, independent researchers seem to be ahead of the government. Fedasiuk, the CSET researcher who uncovered new trends in the Chinese military's use of AI, notes that the Commerce Department doesn't have the time to go through a dataset of thousands of government procurement documents — like he did for one year— to understand the PLA's priorities, how they are going about securing critical technology, and even how some U.S. AI technology is ending up in their hands. "Some members of Congress have been frustrated with the Department of Commerce and the Bureau of Industry and Security,"

Fedasiuk says, referring to the government organ tasked with preventing China from obtaining U.S. advanced technology. "But it is a manpower and language issue – very few staff speak Chinese, and it can be a challenge to distinguish between similarly-named companies, or to understand their ownership structure."

In addition to helping the U.S. government, however, <u>Matt Korda</u>, a researcher at the Federation of American Scientists, says independent open source researchers can help keep the government accountable — and keep the public informed.



Dome structures covering silo construction sites identified by Matt Korda, a researcher at the Federation of American Scientists. Credit: Federation of American Scientists

Last summer, from his DC office, Korda painstakingly scanned satellite imagery of the Chinese desert just below the Mongolian border, looking for nuclear silos. After walking across the map with his cursor, he found <u>silo fields</u> that indicated a significant expansion of China's nuclear arsenal and <u>set off</u> alarm bells among international policy makers about China's nuclear capabilities.

Korda says the government's response to the discovery was muted, indicating that U.S. officials may have already been aware of the silo field and decided not to publicly disclose it. Up to that point, Korda says, the U.S. government was making statements about advanced Chinese nuclear capabilities that didn't match up with what the public knew.

"We saw what happened when they said 'just trust us' during the Iraq war," he says. "Now, open source researchers are able to confirm or push back. We are filling in the gaps."

Indeed, with a new generation of China researchers obsessing over these gaps and holding the government and each other accountable, China may not end up being the black box many people fear it to be. Xiao, for instance, calls open-source research a "new discipline" and hopes to see more researchers taking advantage of these inventive methods over the next decade.

• "We saw what happened when they said 'just trust us' during the Iraq war...now, open source researchers are able to confirm or push back."

"

— Matt Korda, Researcher at the Federation of American Scientists

Back in his Minnesota bedroom-cum-office, Zenz certainly remains committed to the cause. "It is completely addictive — my wife can testify to this," he says unapologetically. "It's a trait and a fault."

As he describes the urgency of his Xinjiang work and therefore the difficulty he has setting boundaries around it and his personal life, a large sign on the dresser practically screams for the room's inhabitants to "RELAX." But Zenz is too excited — both about the data and about what he can uncover with it.

"I was very surprised in the beginning, and I continue to be surprised, especially now, about how much is available," he says. "The flexibility of a single mind can penetrate China's publicly available systems and elephantine bureaucracy."



Katrina Northrop is a journalist based in Washington D.C. Her work has been published in *The New York Times, The Atlantic, The Providence Journal,* and *SupChina.* <u>@NorthropKatrina</u>

COVER STORY



Pole Position

BY EYCK FREYMANN

In public, Chinese diplomats and climate negotiators deny that they see any link between climate change and geopolitics. But there is a deeply cynical consensus within China's academic and policy communities that climate change creates geopolitical opportunities that China can exploit — and must exploit before its rivals do. Greenland was the proof of concept for this strategy. And it caught the U.S. flat-footed.

THE BIG PICTURE



🗕 Q & A



Jörg Wuttke on China's Self-Destruction BY ANDREW PEAPLE

Transsion's Triumph by garrett o'brien

growth potential.

A look at Transsion's monumental growth, unique marketing strategies and future

The EU Chamber of Commerce in China president talks about China's self-inflicted problems; how he gets away with being so outspoken; and why he believes in China's comeback gene.



Visit News Products Store

News Products

Our best open-source research on Chinese companies, as well as industry guides to 100 of the most influential people in a China-focused industry.

