

Securing Software by Construction

Jean Yang

Harvard Medical School/
Carnegie Mellon University

April 11, 2016

Our Lives Run on Software

KIM ZETTER SECURITY 08.04.15 7:00 AM

HACKERS CAN SEIZE CONTROL OF ELECTRIC SKATEBOARDS AND TOSS RIDERS

ANDY GREENBERG SECURITY 09.10.15 7:00 AM

GM TOOK 5 YEARS TO FIX A FULL-TAKEOVER HACK IN MILLIONS OF ONSTAR CARS

ANDY GREENBERG SECURITY 07.29.15 7:00 AM

HACKERS CAN DISABLE A SNIPER RIFLE—OR CHANGE ITS TARGET

ANDY GREENBERG SECURITY 07.21.15 6:00 AM

Smart homes

HACKERS REMOTELY KILL A JEEP ON THE HIGHWAY—WITH ME IN IT

But first we need to “solve” security!

State of the Art

Academia



Encrypted
databases



Undo
mechanisms

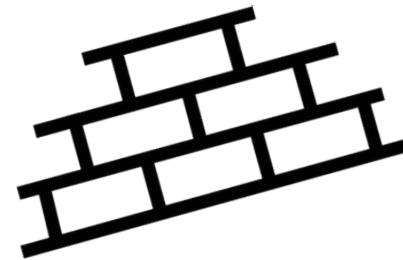


Program
analyses



Provably secure
software

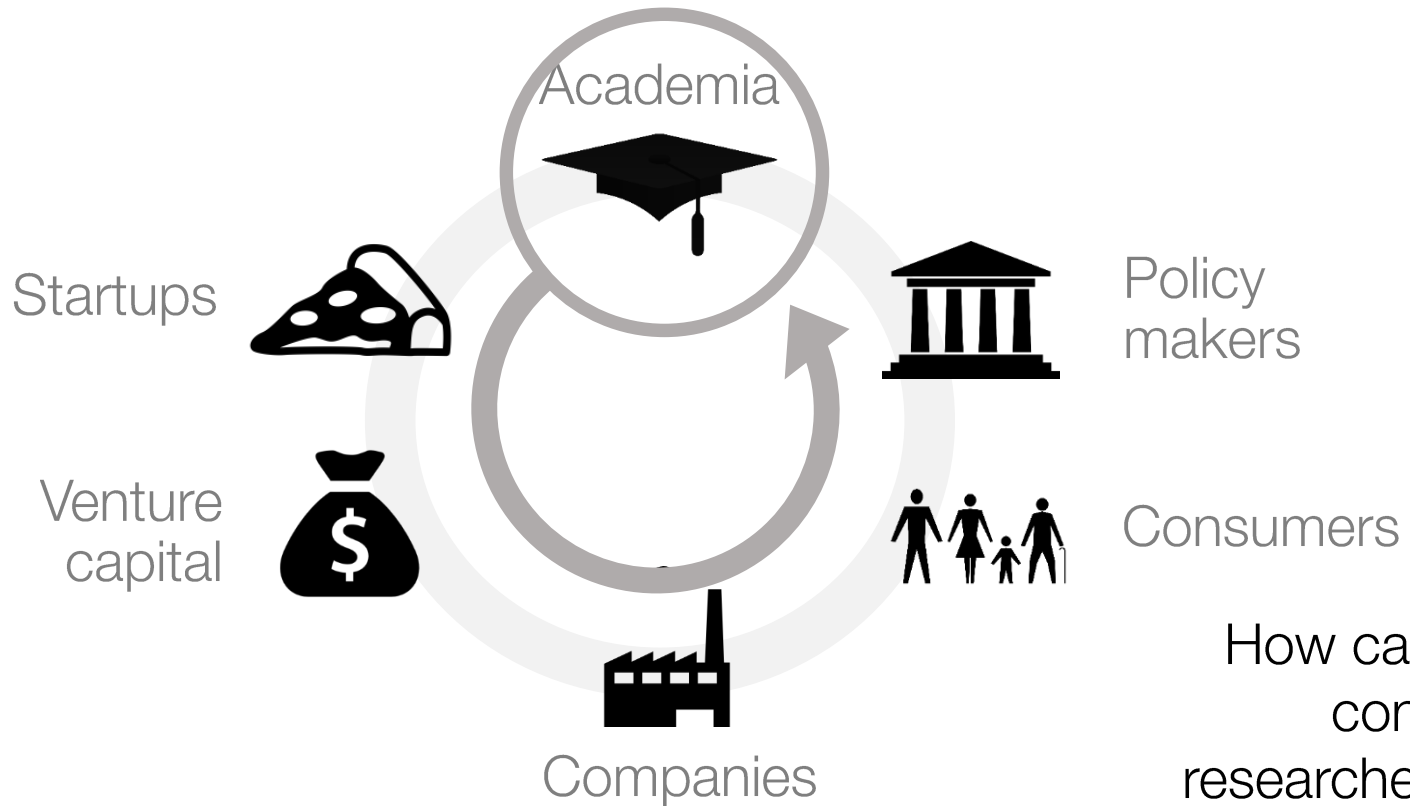
Industry



Firewalls

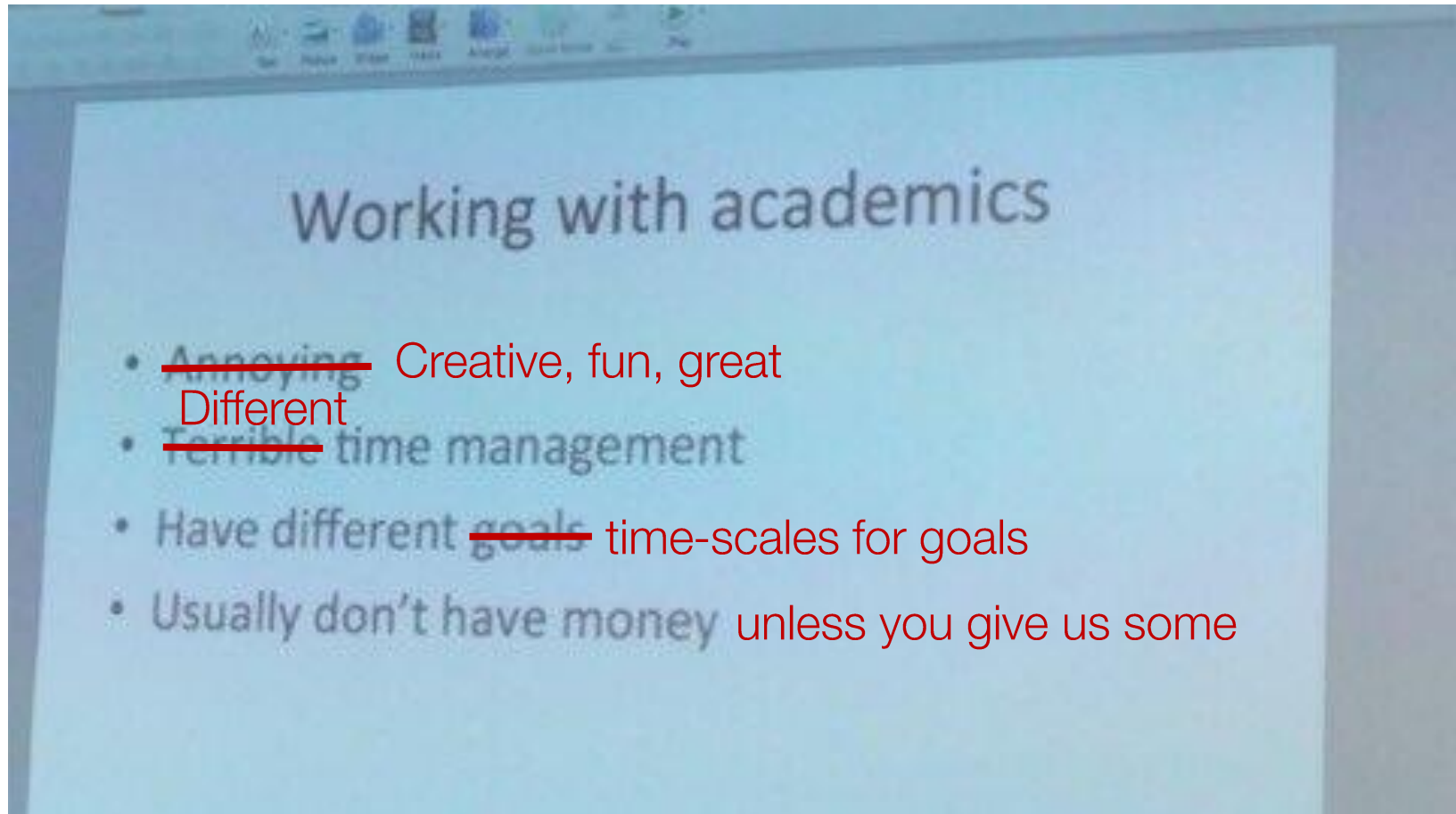
The big question:
How can we take advantage
of research ideas in practice?

This Talk



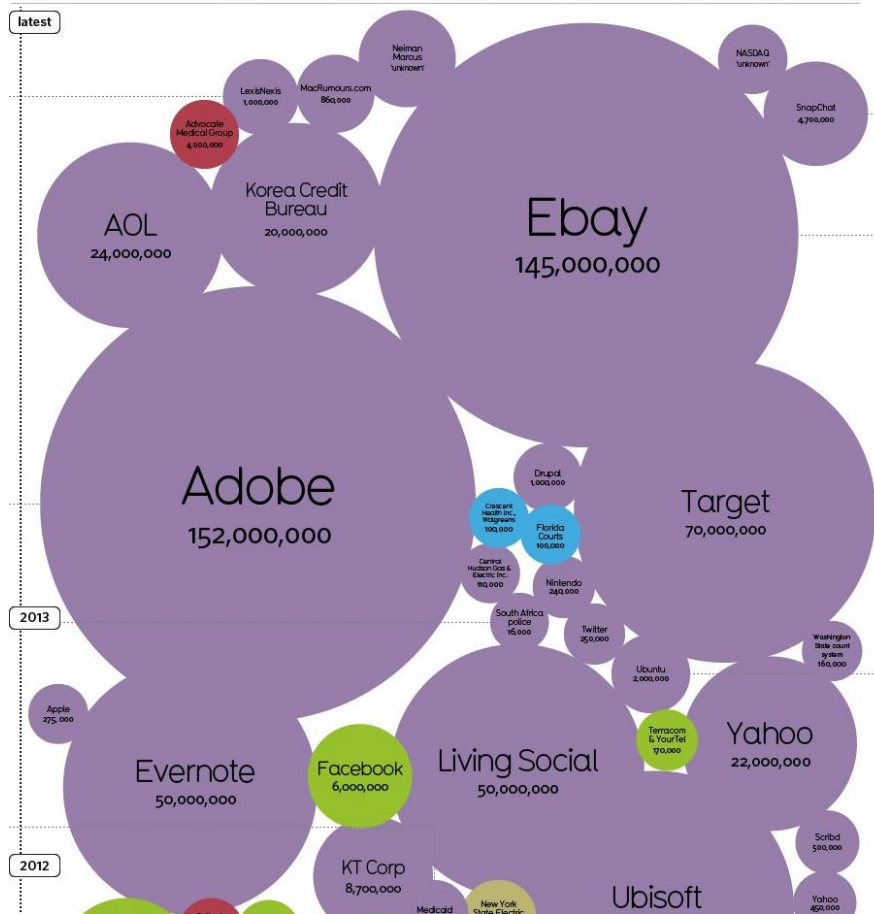
How can we
connect
researchers to
everyone else?

Secondary Goals of Talk



Part I: What Do Academics Think About?

Problem I'm Solving: Protecting Sensitive Data is Hard



- Nobody is surprised to hear about data breaches.
- Reasoning about code is difficult to scale.
- Left with heuristics and little hope about information security.



Why Aren't Existing Approaches Enough?

Encrypting Data



But people often *are* protecting data—though incorrectly.

Reactive Security



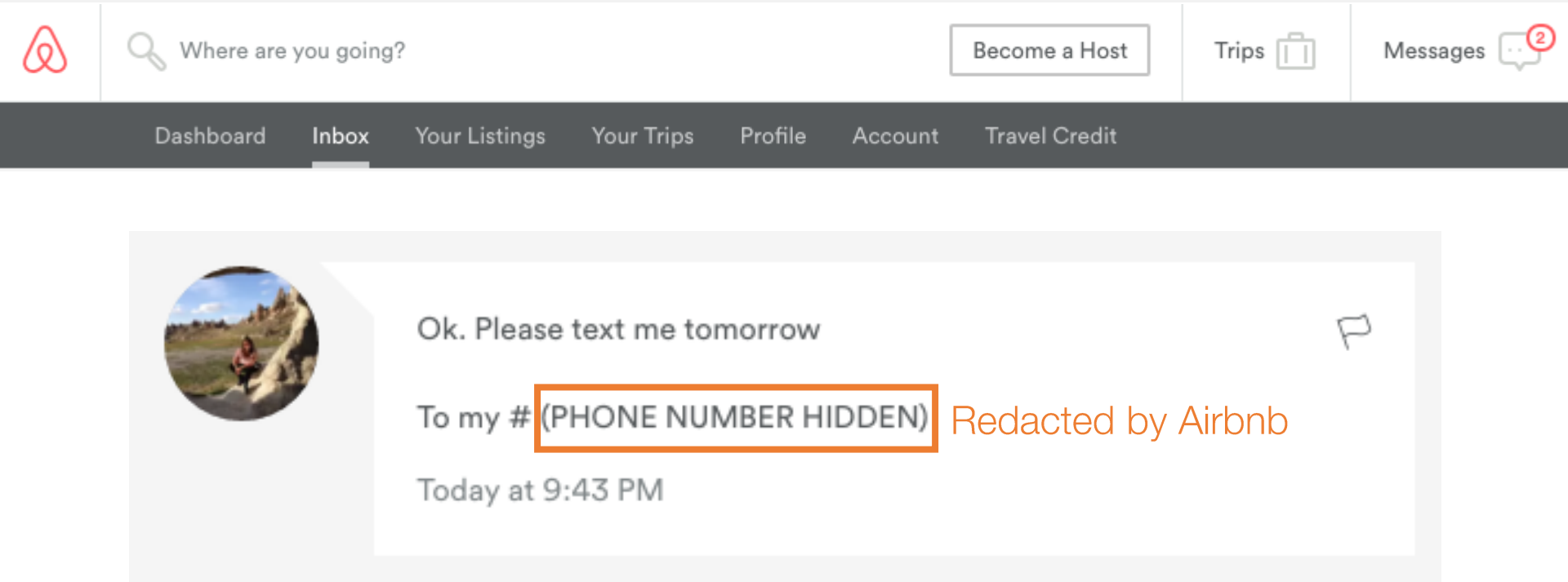
But leaves system builders a step behind.

My Approach

Factor out policy checks to reduce opportunity for leaks.

- Programmer specifies high-level policies about how sensitive data may be used.
- Rest of program is policy-agnostic.
- System manages policies automatically.

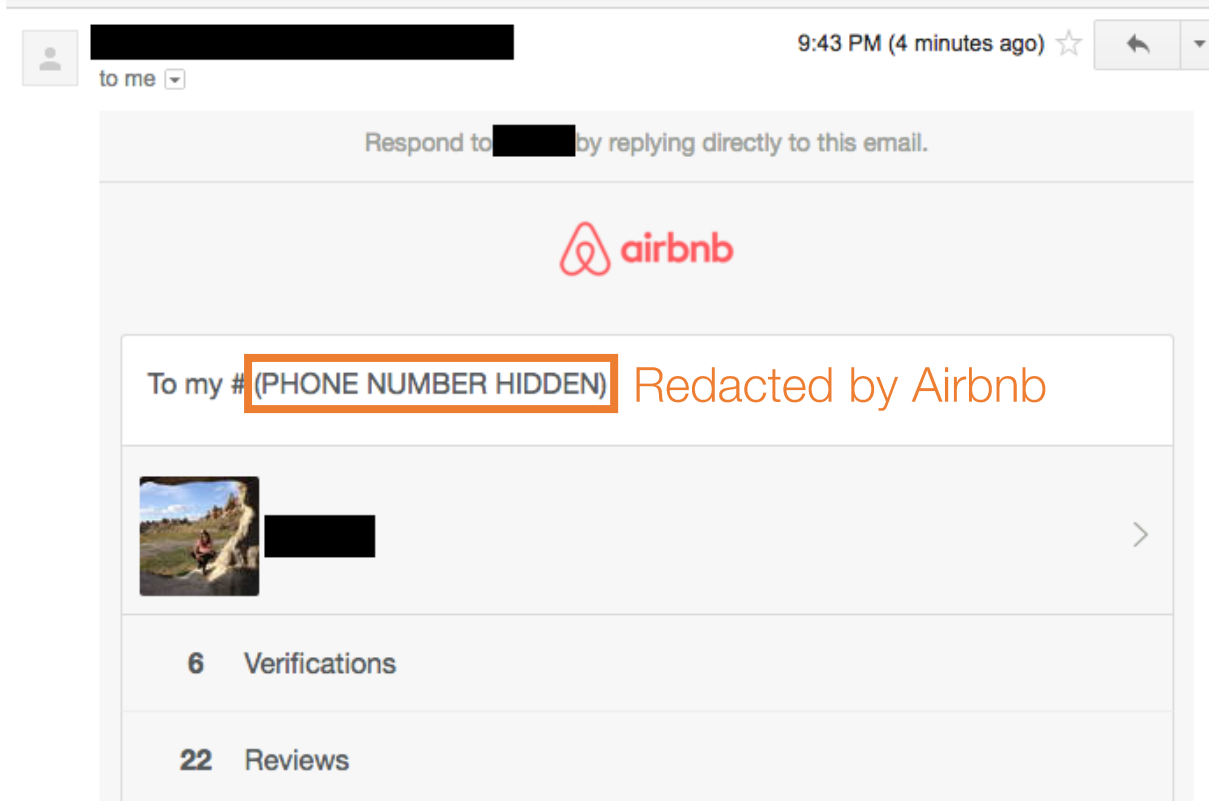
Goal: Keep Sensitive Information Private



Airbnb has a policy of blocking phone numbers so communications happen through their application.

Example courtesy of Chelsea Voss

Need to Make Sure Information Protected in All Views



Phone number remains redacted in email view.

Example courtesy of Chelsea Voss

Missed a spot!

The screenshot shows the Airbnb user interface. At the top, there is a search bar with the text "Where are you going?", a "Become a Host" button, and navigation links for "Trips", "Messages" (with a notification badge of 2), and "Help". Below this is a dark navigation bar with links for "Dashboard", "Inbox", "Your Listings", "Your Trips", "Profile", "Account", and "Travel Credit". A light blue promotional banner for "Earn \$100 travel credit" is visible, with buttons for "Invite Friends" and "Later". Below the banner is a message preview for "All Messages (2)". The message is from a user with a profile picture of a horse, sent at 9:43 PM. The message content is "To my # [redacted] Washington, DC (Mar 15 - 16, 2016)". The status is "Accepted" with a value of "\$115". A red arrow points to the redacted phone number in the message preview, with a red text annotation: "Actual phone number! Redacted by me and not Airbnb."

Phone number is visible in message preview.

Example courtesy of Chelsea Voss

Programmers Must Navigate “Policy Spaghetti”

```
if ($stype === "p")
    $stype = "r";
if ($this->privChair && !$stype && $Conf->timeUpdatePaper())
    $this->limitName = "all";
else if (($me->privChair && $stype == "act")
    || ($me->isPC
    && (!$stype || $stype == "act" || $stype == "all")
    && $Conf->can_pc_see_all_submissions()))
    $this->limitName = "act";
else if ($me->privChair && $stype == "unm")
    $this->limitName = "unm";
else if ($me->isPC && (!$stype || $stype == "a" || $stype == "unm"))
    $this->limitName = "s";
else if ($me->isPC && ($stype == "und" || $stype == "undec"))
    $this->limitName = "und";
else if ($me->isPC && ($stype == "acc" || $stype == "revs"
    || $stype == "reqrevs" || $stype == "req"
    || $stype == "lead" || $stype == "rable"
    || $stype == "manager"))
    $this->limitName = $stype;
else if ($this->privChair && ($stype == "all" || $stype == "unsub"))
    $this->limitName = $stype;
else if ($stype == "r" || $stype == "rout" || $stype == "a")
    $this->limitName = $stype;
else if ($stype == "rable")
    $this->limitName = "r";
else if (!$me->is_reviewer())
    $this->limitName = "a";
else if (!$me->is_author())
    $this->limitName = "r";
else
    $this->limitName = "ar";

// track other information
$this->allowAuthor = false;
if ($me->privChair || $me->is_author()
    || ($this->amPC && $Conf->submission_blindness() != Conference::BLIND_ALWAYS))
```

396,16 11%

```
// if a complex request, or a form upload, don't search
foreach ($_REQUEST as $k => $v)
    if ($k != "p" && $k != "paperId" && $k != "m" && $k != "mode"
        && $k != "forceShow" && $k != "go" && $k != "actas"
        && $k != "ls" && $k != "t"
        && !isset($_COOKIE[$k]))
        return false;

// if no paper ID set, find one
if (!isset($_REQUEST["paperId"])) {
    $q = "select min(Paper.paperId) From Paper ";
    if ($me->isPC)
        $q .= "where timeSubmitted>0";
    else if ($me->has_review())
        $q .= "join PaperReview on (PaperReview.paperId=Paper.paperId and PaperReview.contactId=$me->contactId)";
    else
        $q .= "join ContactInfo on (ContactInfo.paperId=$me->paperId and ContactInfo.contactId=$me->contactId and ContactInfo.conflictType != CONFLICT_AUTHOR . ")";
    $result = $Conf->q($q);
    if (($paperId = edb_row($result)))
        $_REQUEST["paperId"] = $paperId;
    return false;
}

// if invalid contact, don't search
if ($me->is_empty())
    return false;

// actually try to search
if ($_REQUEST["paperId"] != "p")
    $_REQUEST["paperId"] = "p";
$search = new PaperSearch($me, array("q" => $_REQUEST["paperId"], "t" =>
    defval($_REQUEST, "t", 0)));
$pl = $search->paperList();
if (count($pl) == 1) {
    $pl = $search->session_list_object();
```

2143,24 93%

Code from
HotCRP
conference
management
system

Highlighted: conditional permissions checks everywhere.

Solution: Allow Programmers to Attach Policies Directly to Data

The language and runtime manage policies so the programmer does not need to.



Policy-agnostic programming model and guarantees
[POPL '12]



Improved semantics based on multi-execution
[PLAS '13]



Extending programming model across database
[PLDI '16]

Policy-Agnostic Programming Factors Out Policies

```
class Event(Model):
    VISIBILITY = (('E', 'Everyone'), ('G', 'Guests' ))

    name = CharField(max_length=256)
    location = CharField(max_length=512)
    time = DateTimeField()
    description = CharField(max_length=1024)
    visibility = CharField(max_length=1, choices=VISIBILITY, default='E')

    @jeeves
    def has_host(self, host):
        return EventHost.objects.get(event=self, host=host) != None

    @jeeves
    def has_guest(self, guest):
        return EventGuest.objects.get(event=self, guest=guest) != None

    @staticmethod
    def jeeves_get_private_name(event):
        return "Private event"

    @staticmethod
    def jeeves_get_private_location(event):
        return "Undisclosed location"

    @staticmethod
    def jeeves_get_private_time(event):
        return datetime.now(tz=pytz.utc)

    @staticmethod
    def jeeves_get_private_description(event):
        return "An event."

    @staticmethod
    @label_for('name', 'location', 'time', 'description')
    @jeeves
    def jeeves_restrict_event(event, ctxt):
        if event.visibility == 'G':
            return event.has_host(ctxt) or event.has_guest(ctxt)
        else:
            return True
```

Model

```
<h2 class="form-heading">Edit Your Event.</h2>
<div class="form-group">
<label for="name" class="control-label"><a rel="tooltip" title="T000."><span class="glyph">
<div class="row">
<input type="text" class="form-control" name="name" id="name" value="{{ concretize(name)}}">
</div>
</div>
<div class="form-group">
<label for="location" class="control-label"><a rel="tooltip" title="T000."><span class="glyph">
<div class="row">
<input type="text" class="form-control" name="location" id="location" value="{{ concretize(l">
</div>
</div>
<div class="form-group">
<label for="time" class="control-label"><a rel="tooltip" title="T000."><span class="glyph">
<div class="row">
<input type="datetime" class="form-control" name="time" id="time" value="{{ concretize(time)">
</div>
</div>
</div>
<div class="form-group">
<label for="description" class="control-label"><a rel="tooltip" title="T000."><span class="glyph">
<div class="row">
<div class="col-xs-4">
<textarea class="form-control" name="description" id="description">{{ concretize(description)">
</div>
</div>
</div>
</div>
<div class="form-group">
<label for="visibility" class="control-label"><a rel="tooltip" title="T000."><span class="glyph">
<div class="row">
<div class="col-xs-4">
<div class="btn-group">
<button type="button" class="btn btn-default {% if visibility=='E' %}active{% endif %}" dat
<button type="button" class="btn btn-default {% if visibility=='G' %}active{% endif %}" da
</div>
</div>
</div>
</div>
<button class="btn btn-primary" type="submit" value="Submit" onClick="myApp.showPleaseWait();">Submi
</form>
```

```
@login_required
@request_wrapper
@jeeves
def jeeves_profile_view(request, user_profile):
    profile = UserProfile.objects.get(username=request.user.username)
    if profile == None:
        profile = user_profile

    if request.method == 'POST':
        user_profile.username = request.user.username
        profile.name = request.POST.get('name', '')
        profile.email = request.POST.get('email', '')
        profile.save()

    host_events = EventHost.objects.filter(event=profile).all()
    guest_events = EventGuest.objects.filter(event=profile).all()

    return ("profile.html", {
        "profile": profile,
        "is_own_profile": request.user.username==user_profile.username,
        "host_events": host_events,
        "guest_events": guest_events,
        "which_page": "profile",
    })

def register_account(request):
    if request.user.is_authenticated():
        return HttpResponseRedirect("index")

    if request.method == 'POST':
        form = UserCreationForm(request.POST)
        if form.is_valid():
            user = form.save()
            user.save()

            UserProfile.objects.create(
                username=user.username,
                email=request.POST.get('email', ''),
            )
```

- Centralized policies.
- Policy-agnostic program.
- Runtime differentiates behavior.

View

Controller

Jeeves Language and Execution Model

① Runtime propagates values and policies.

x = 0

if

actual number	HIDDEN
---------------	--------

== "867-5309":

x += 1

return x

② Runtime solves for values to show based on policies and viewer.

print {  } 

1

 **print** {  }

0

Semantics of Output

Statement
evaluation

$$\Sigma, S \Downarrow V_p, oc: R$$

Expression
evaluation

$$\Sigma, E \Downarrow_{pc} \Sigma', V$$

$$\begin{aligned} \Sigma, E_{oc} &\Downarrow_{\emptyset} \Sigma_{oc}, V_{oc} \\ \Sigma_{oc}, E_r &\Downarrow_{\emptyset} \Sigma_r, V_r \end{aligned}$$

Evaluate output context
and expression to print.

$$\{k_1 \dots k_n\} = \text{closeK}(\text{labels}(E_{oc}) \cup \text{labels}(E_r), \Sigma_2)$$

$$E_p = \lambda x. \text{true} \wedge_f \dots \wedge_f \Sigma_2(k_n)$$

Retrieve labels
and policies.

$$\Sigma_r, (E_p V_{oc}) \Downarrow_{\emptyset} \Sigma_p, V_p$$

Evaluate policies applied
to the output context.

pick pc such that $pc(V_{oc}) = oc, pc(V_r) = R, pc(V_p) = \text{true}$

$$\Sigma, \text{print} \{E_{oc}\} E_r \Downarrow V_p, oc: R$$

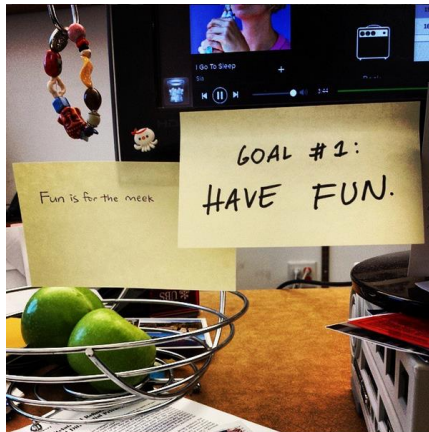
Policies




Output
context

Result

Defacet using
satisfying policy
assignment.

The Pain of Production-Testing a Research Prototype



 [Redacted] via mit.edu 3/11/12 ☆  

to Jean ▾

Hi Jean,

I don't know what the problem is, but I cannot log in to the submission system now. It keeps loading but the page just doesn't show up. Is there any problem with the **server**?

Thanks,
[Redacted]

 [Redacted] 3/12/12 ☆  

to Jean ▾

Hi,

I hate to tell you this, but the **server** seems **down** again...

Cheers,
[Redacted]

 Armando Solar-Lezama [Redacted] 3/13/12 ☆  

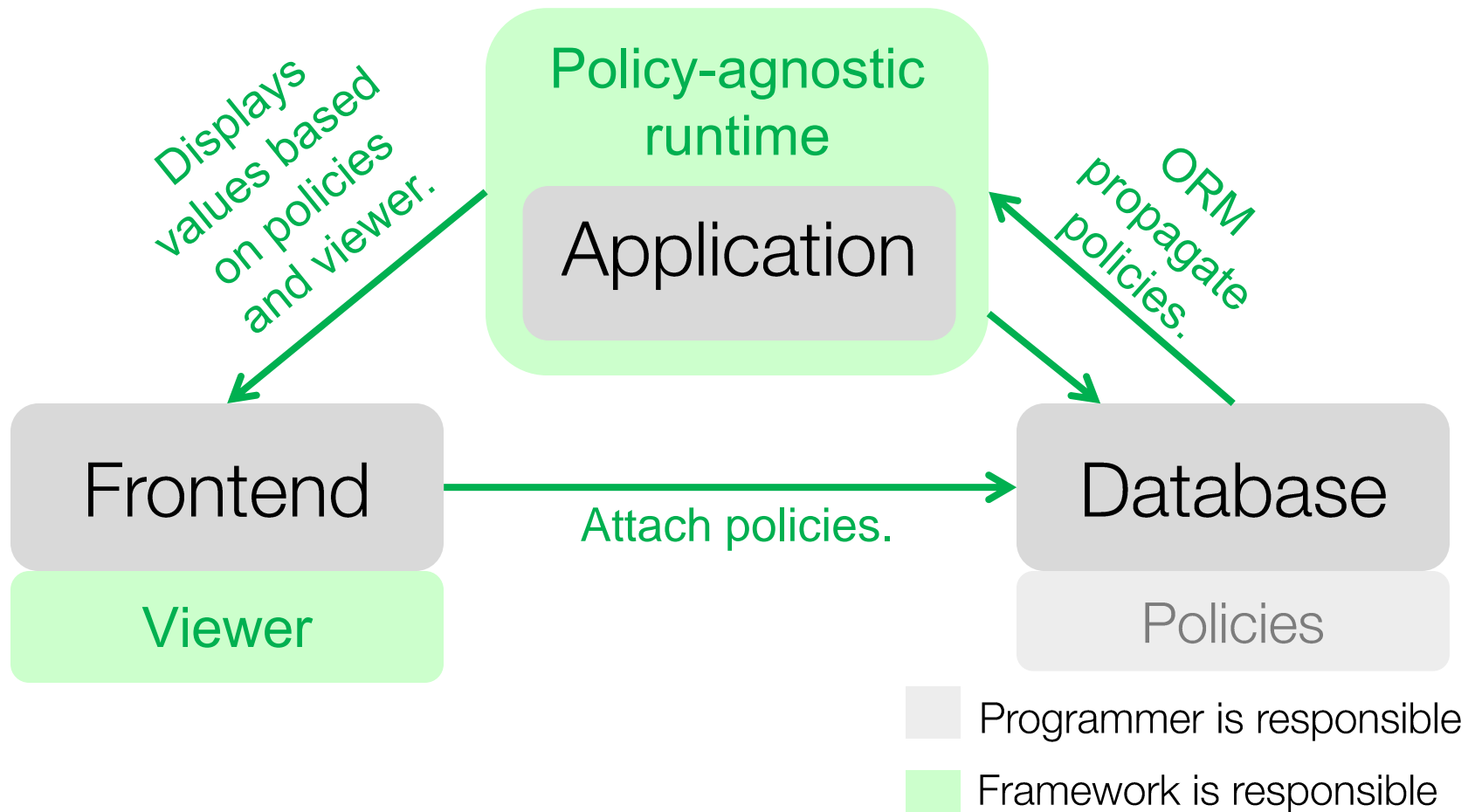
to Jean, Kuat ▾

Prudence is for the meek. Let's do whatever it takes to get **the** system to stay up. Are you available to talk on the phone now?

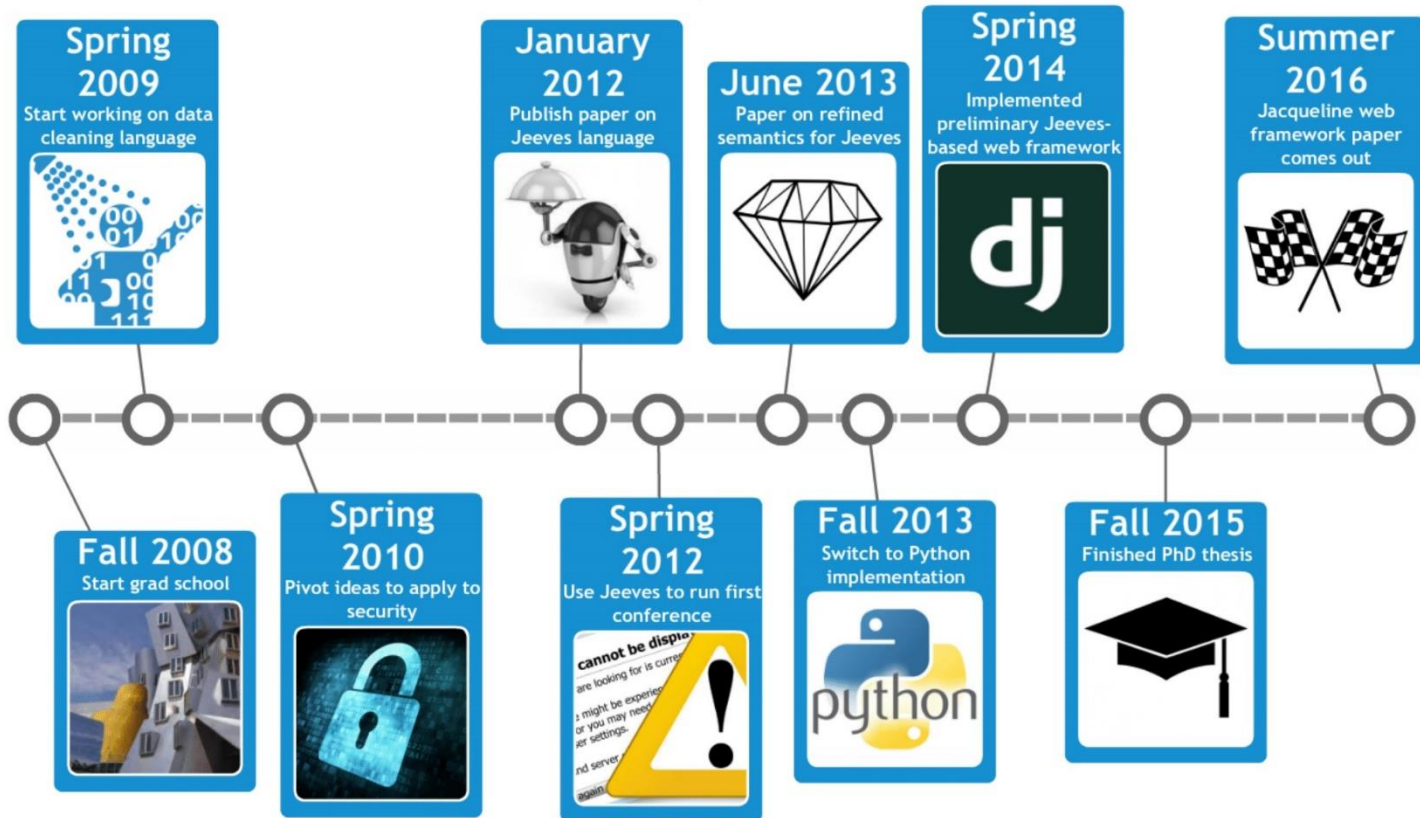
Lessons Learned

- Need a solution for running out of memory.
- Need a story for extending language-level guarantees to the database.
- But, in good news, web programs are often short and simple.

Jacqueline, a Policy-Agnostic Web Framework

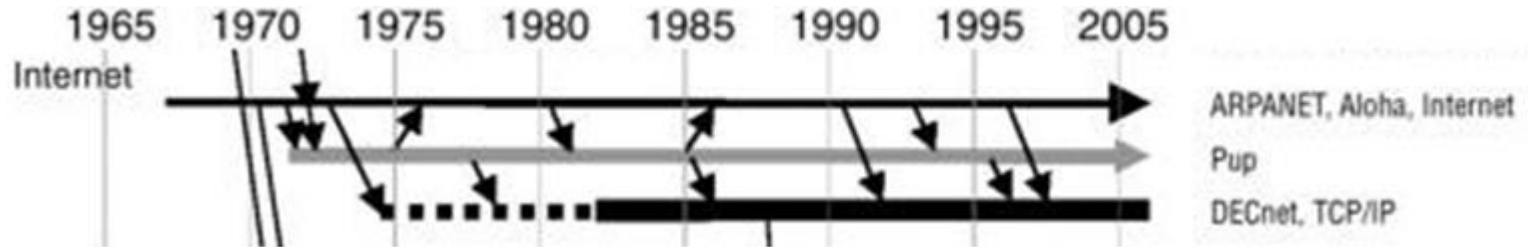


Research is Slow



At this point, we have proposed a new programming model and de-risked the problem for people in industry.

Be Patient with Us!



- Research takes time.
- Adoption into the mainstream can take even more time.
- Many features in modern programming languages were incubated in research decades ago!

Part II:
How Can We Use
Research Results in
the Real World?

Barriers to Industry Adoption

In large companies:

- **Managers** need to fight status quo.
- **Programmers** need to manage legacy code.



Sam Altman ✓

@sama



I would like YC to fund dozens of computer security companies in the next couple of years. Feels like the world is very exposed.

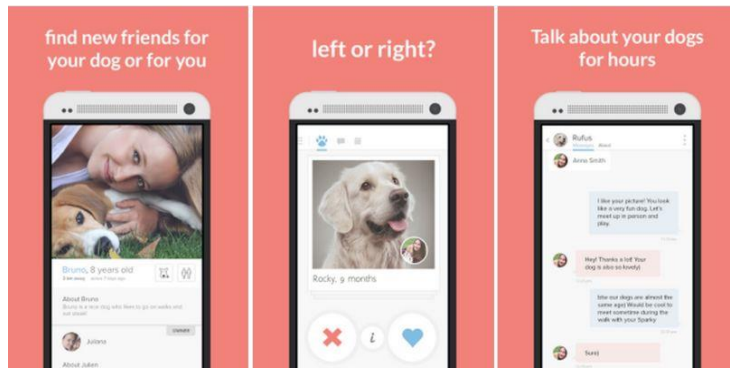
4:46 PM - 8 Jul 2015

↩️ ↻ 198 ★ 346

What about the startup route to tech transfer?

Security is no Tindog

The Hot New Silicon Valley Startup





Fun concept. Slick design.
Toddler nephew can use it.
Integrates with your life.

Startup that Helps Us
Secure Our Software



Technical concept. Verifiable
by experts. Requires
infrastructure change.

Unique Challenges for Security Startups

- Security is expensive. 
- Concept is highly technical.
- No flashy demos.
- Adoption requires client expertise and/or trust. 
- Solving a technical problem != building a product.



Following

"A major reason security products fail is because they're made by security people."
-[@JustinSomaini](#) on the importance of user experience



Justin Somaini,
Chief Security
Officer, SAP

Cybersecurity Factory

HIGHLAND
CAPITAL PARTNERS



An 8-week accelerator that gives teams:



\$20,000



A network



Office space



Legal
support



Maxwell Krohn



Raj Shah



David Ting

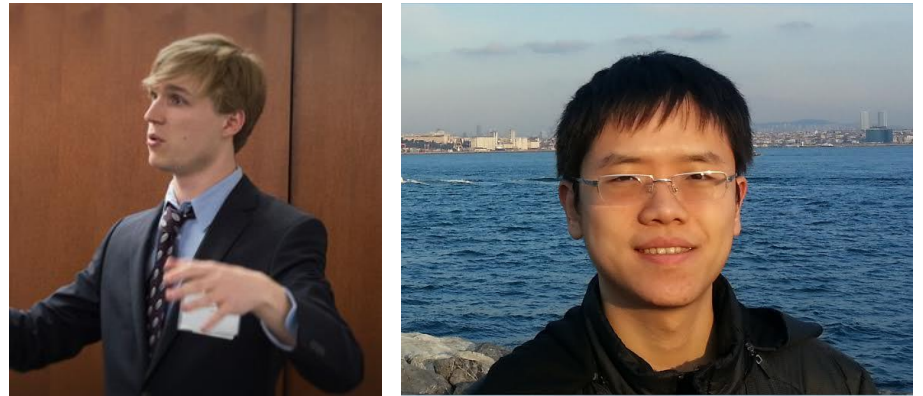


Focused mentorship

Summer 2015 Cohort



Aikicrypt:
Outsourcing data
securely to the
cloud.



Oblivilock:
Protecting data and
metadata in the
cloud.

“I thought it was hard to sell my research. It’s much harder to sell something for money.”

Christopher Fletcher, MIT PhD student,
Cybersecurity Factory participant

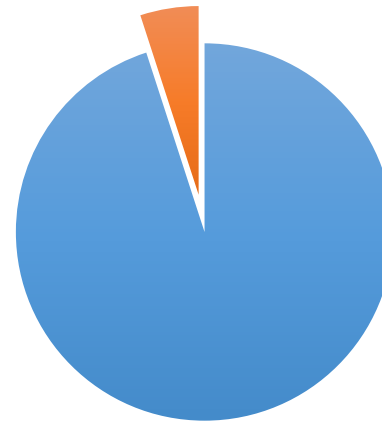
How Teams Spent the Summer

How Teams Thought They Would Spend Time



- Talking to customers and working on pitches
- Coding

How Teams Actually Spent Time



- Talking to customers and working on pitches
- Coding



Andy Fraser
@andy_utoronto



Follow

@jeanqasaur @DrHallba @jmseabrook
evidence for that completely random statement? May be true may not, but random numbers don't equal truth.

12:59 PM - 1 Apr 2016



John Seabrook
@jmseabrook



Following

@andy_utoronto @jeanqasaur @DrHallba It's called "rhetoric" bro.

LIKE

1



6:53 PM - 1 Apr 2016



Biggest Lessons for Teams

- People matter.

Networking can drive innovation.

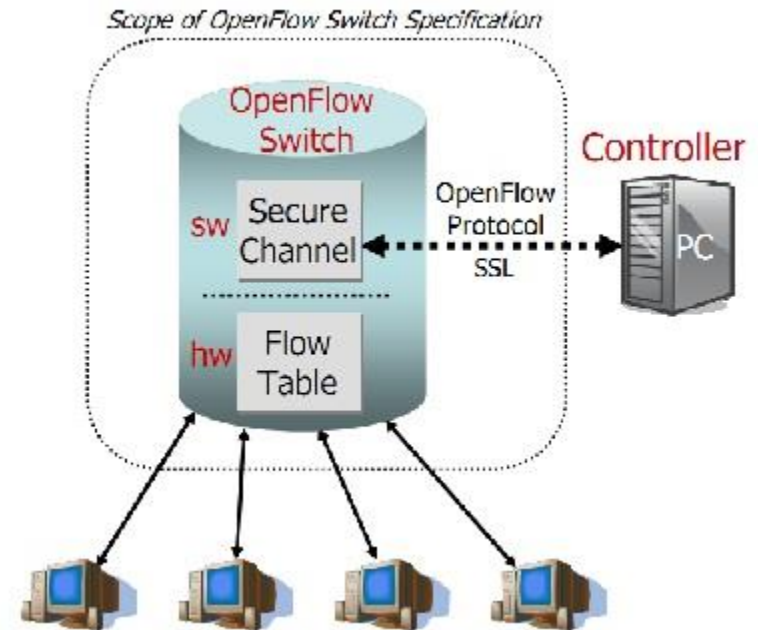
- People matter.

A good product drives conversations.

- People matter.

Finding a target market is crucial.

Fun Discovery: Del Monte Foods is Unexpectedly Hip



Long-Term Goals for Cybersecurity Factory

- Continue running program.
- Commercialize academic security projects.
- Create awareness among investors, clients, and the public.
- More collaboration and partnership with industry.
- Create community of founders interested in technical security problems.

Part III: How To Motivate Customers to Pay for Security?

Insecurity is Expensive

“A report released this month by the Atlantic Council and Zurich Insurance Group estimated that by 2030, an insecure Internet would reduce global economic net benefit by \$90 trillion. In contrast, a completely secure Internet would result in a global net gain of \$190 trillion.”

-Jeff Kosseff, cybersecurity law professor

The Security “Prisoner’s Dilemma”

PRISONER'S DILEMMA

C=cooperate D=defect (don't cooperate)

	C	D
C	-1,-1	-10,0
D	0,-10	-5,-5

YEARS IN PRISON

Lack of individual incentive:

- Requires \$\$.
- Requires more employee training.
- Requires more programmer effort.
- Doesn't currently provide competitive advantage.

We Need to Care More

Consumer Example: Snapchat



Numerous privacy violations,
but valued at \$16 billion with
100 million users.

Policy Example: Dentists



Common to email records
in violation of HIPAA, but
HHS does not audit.

Most Important is Legislative Change

“Intentionally or unintentionally, poorly crafted or outdated laws and technical standards threaten to undermine security, privacy and the viability of our most promising new technologies and networks...” –Joichi Ito

How we can contribute is left as an exercise to the listener.

Conclusion: Many Pieces to Securing Software



But...

If we work together, we can create the right ecosystem to secure our software by construction.

 @jeanqasaur

jeanyang.com

jeeveslang.org

cybersecurityfactory.com