

# Evidence based policy: Why is progress so slow and what can be done about it?

AES17 International  
Evaluation Conference



Canberra 6th September

Nicholas Gruen

[Engruen@lateraleconomics.com.au](mailto:engruen@lateraleconomics.com.au)

[@engruen1](https://twitter.com/engruen1)



Lateral Economics

# Outline

1. Introduction
2. Arteries and capillaries
3. Thick and thin
4. From what to why
5. Evidence based delivery and evaluation
6. Program logic
7. Accountability
8. Institutionalising evidence-based policy



# Introduction



# Evidence-based policy

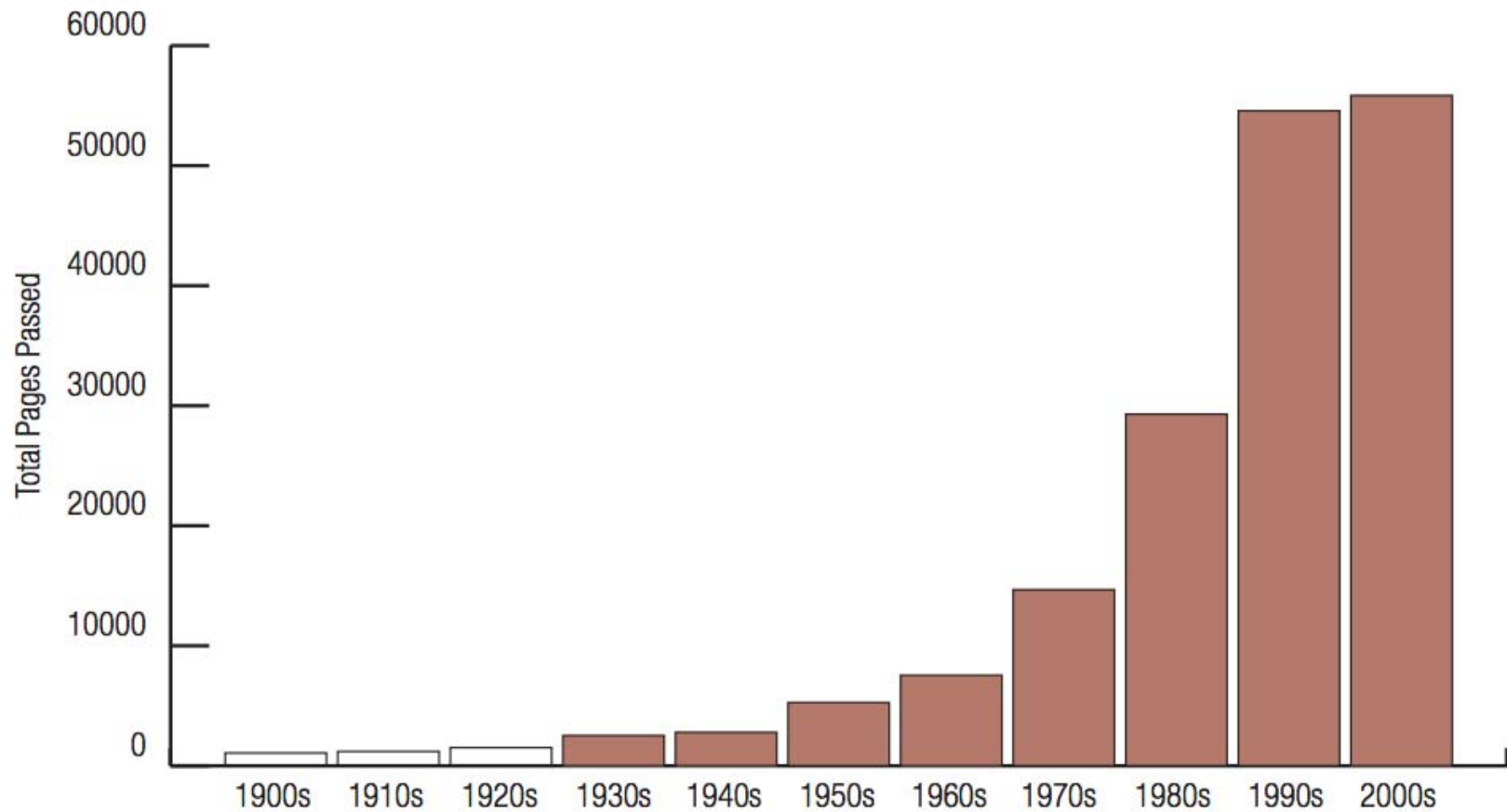
## Conclusion

The NSW Government's program evaluation initiative is largely ineffective, as it is not providing sufficient information to government decision makers on the performance of programs.

No information is provided on the performance of programs that have been evaluated. The information that is provided is limited to a list of programs being evaluated in the upcoming financial year, with little assurance that the right programs are on the list. NSW Treasury and DPC are not using evaluation outcomes to analyse agency funding proposals in their advisory role to the NSW Government.

For program evaluation to be effective, agencies should demonstrate they are evaluating the right programs, and the outcomes from completed evaluations should inform advice to the NSW Government on investment decisions.

### Estimated growth in pages of Australian Government primary legislation

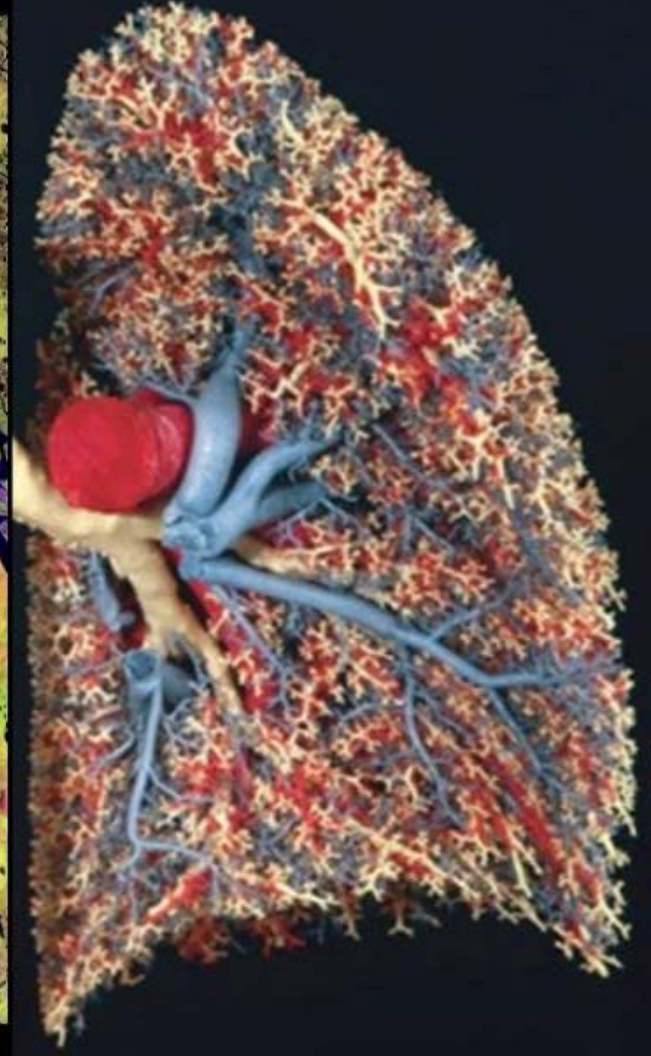


Source: Business Council of Australia, sub. 109, p. 3

# Arteries and capillaries







@ngruen1



Lateral Economics





@ngruen1



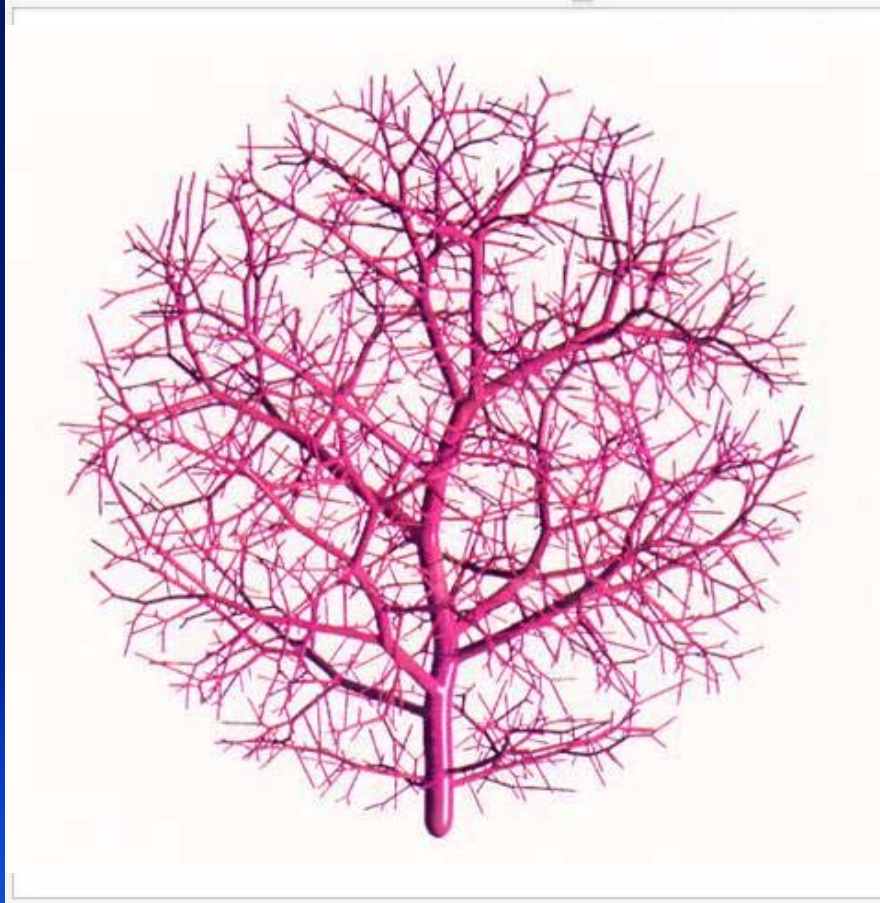
Lateral Economics



# The human world is ...

A hierarchy of trunks and branches  
Arteries and capillaries



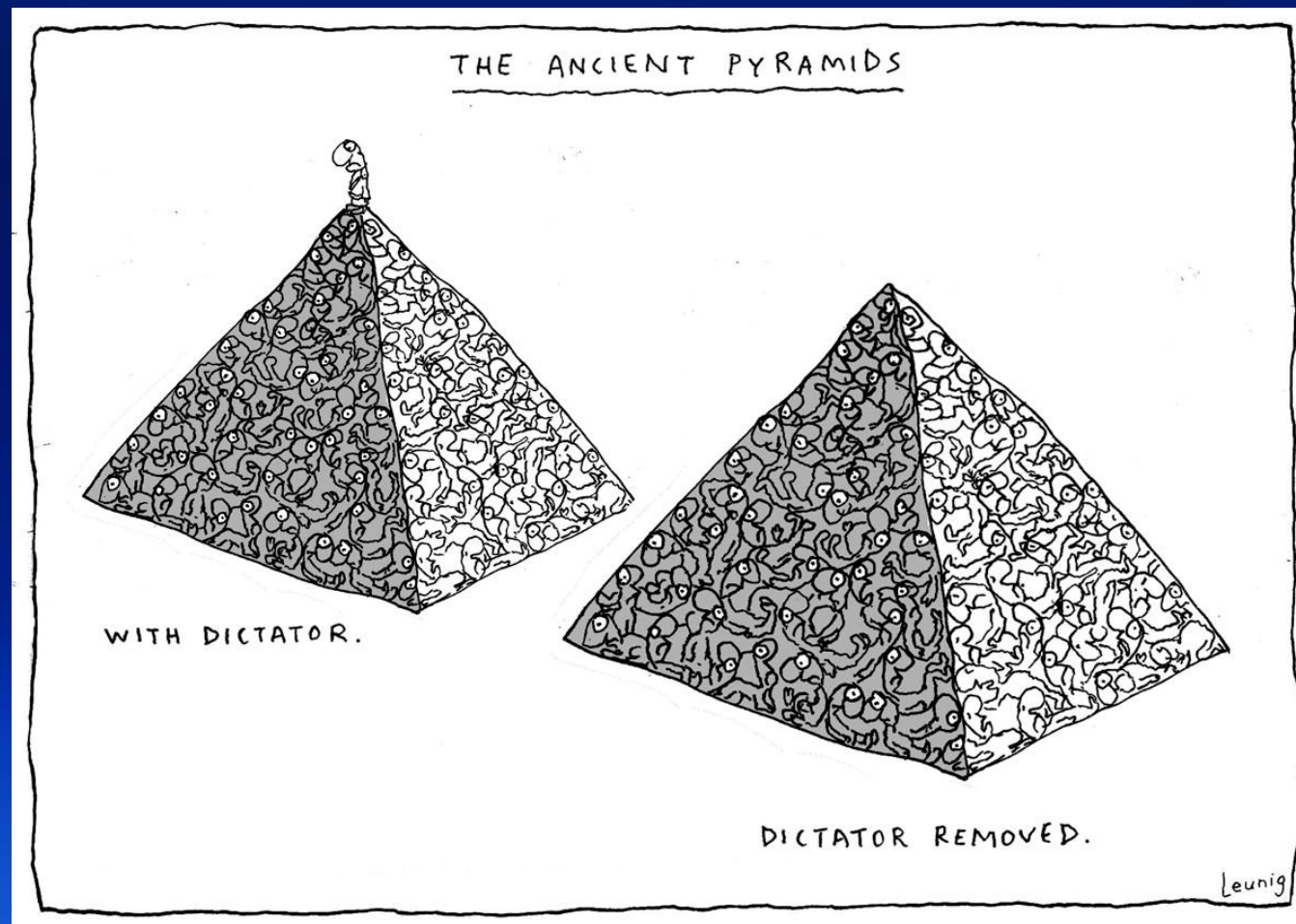


This could be a

- Software program
- Profession or discipline
- Industry
- Encyclopedia
- Catalogue
- Org chart
- Corporate accounts
- Corporate KPIs

# Democracies are hierarchies







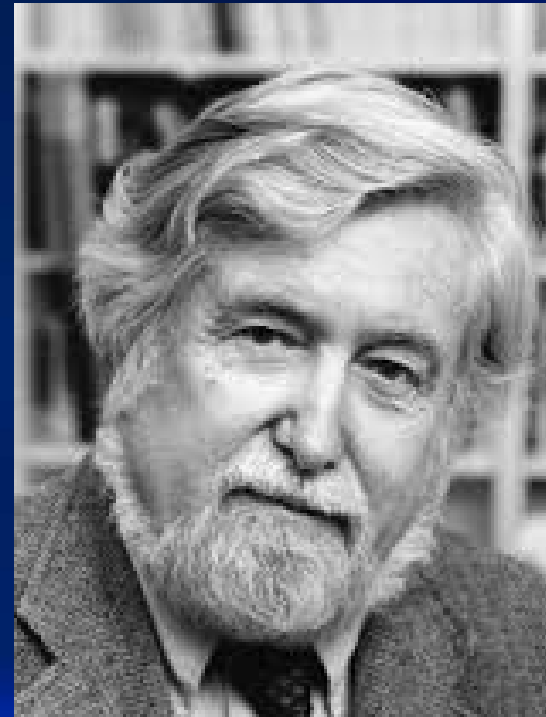
# Thick and thin



# Thick and thin



Gilbert Ryle



Clifford Geertz

# Thick, thin and the understanding of context

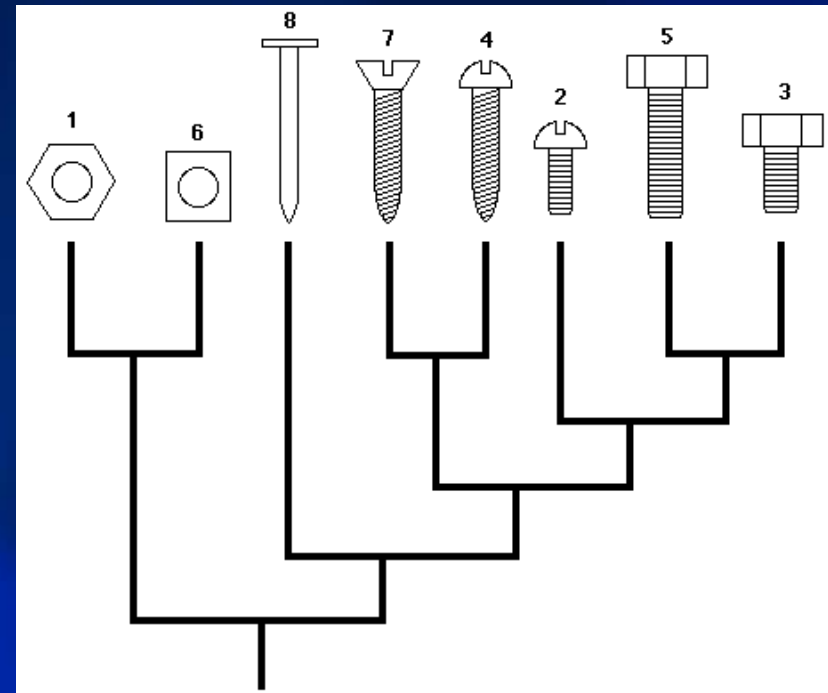
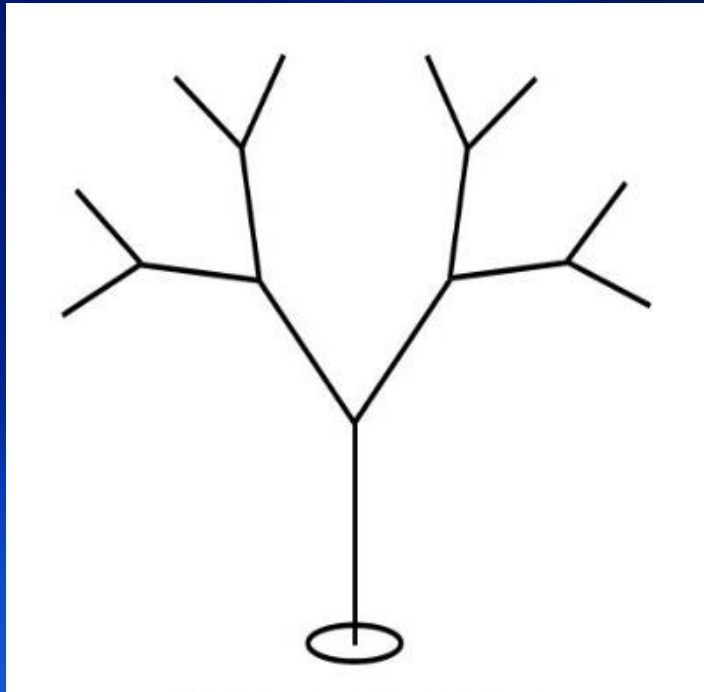


# Thin to thick From what to why

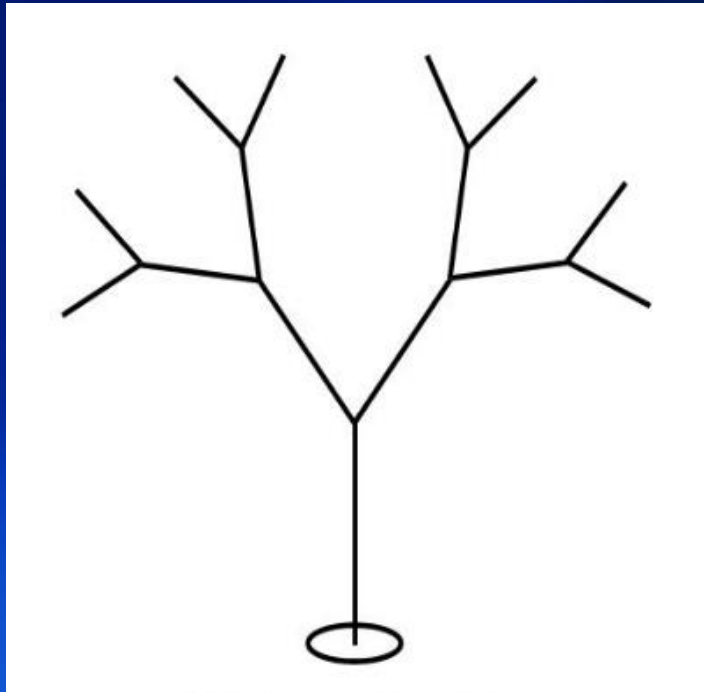




# Thin problems are mechanical



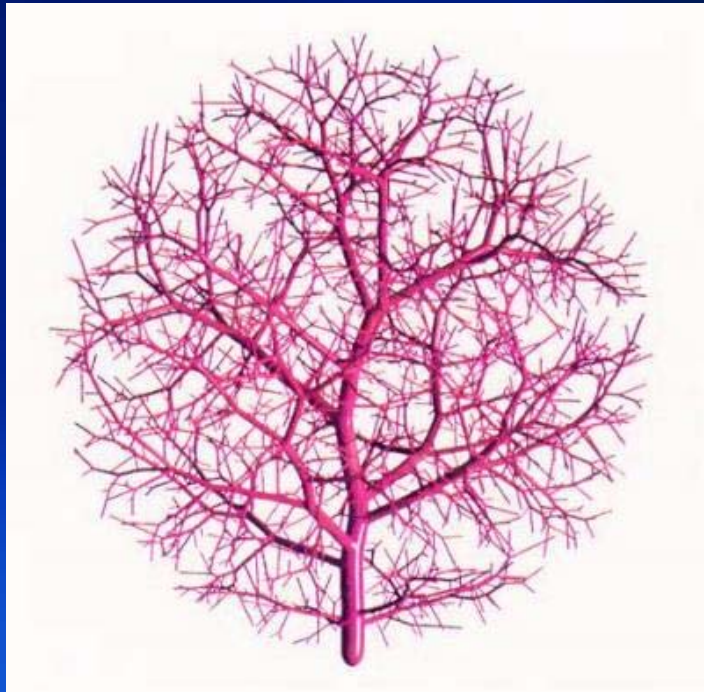
# Thin policy delivery



Top-down policy can work

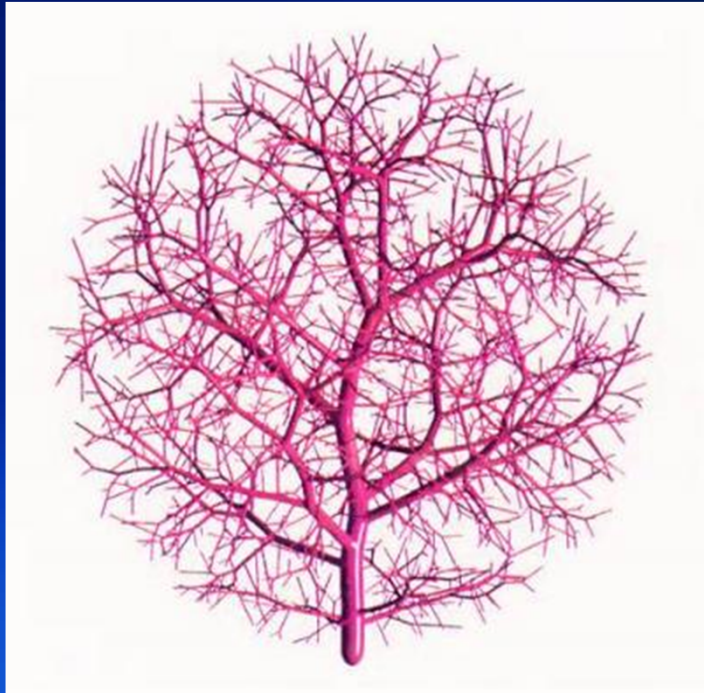
- Tax and family benefits changes
- HECS
- Child Support Agency
- Stroke of the pen deregulation
  - shopping hours, two airline policy

# Thick policy delivery



- Regulation
- IT
- Education
- Health
- Defence
- Transport
- Employment services
- Social support (Indigenous and other)

# Ideologies are thin, issues are thick



Income management

User Choice

Diversity

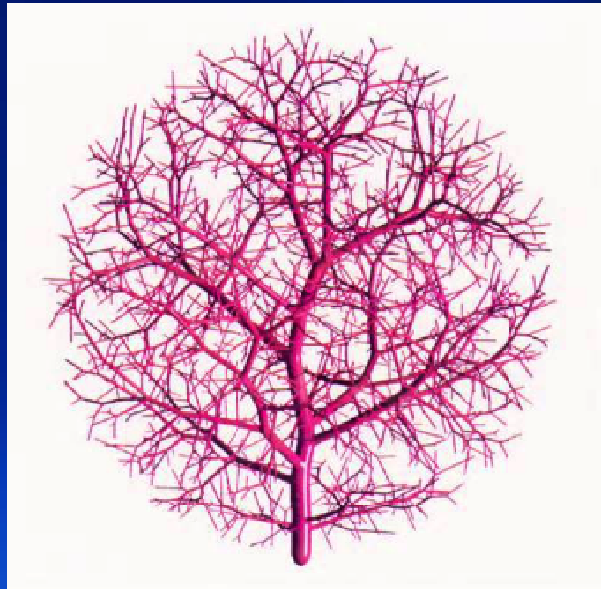
Core values

Individual responsibility

Collective responsibility



# Arteries, capillaries and status



If there were a single cultural predilection in the APS I'd change, it would be the unspoken belief that the development of government policy is a higher function – more prestigious, more influential, more exciting – than delivering results.

Peter Shergold, 2005.

The arteries are willing,  
but the capillaries are weak





# The cult of announceables

@ngruen1



Lateral Economics

23



## Further Education



**48 Secretaries of State**

Responsible for Further Education since 1981

**13,000 qualifications**

In the Further Education space, as noted by Philip Hammond in Budget 2017

## Regional government



**17 years**

Government Offices, 1994-2011

**12 years**

Regional Assemblies, 1998-2010

**1 year**

Local Authority Leaders' Boards, 2009-10

**1992**

'New' unitary authorities created

**2009**

'New' unitary authorities created

Source: Phil Swann, 2016

## Industrial strategy



**2008**

Mandelson's 'new activism'

**2012**

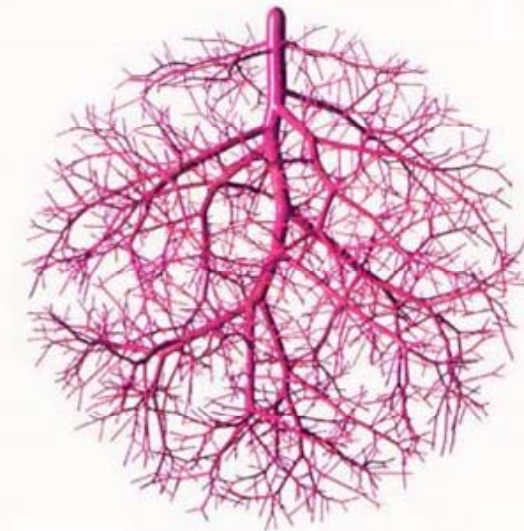
Cable's industrial strategy

**2015**

Javid's 'industrial approach'

**2016**

May's industrial strategy

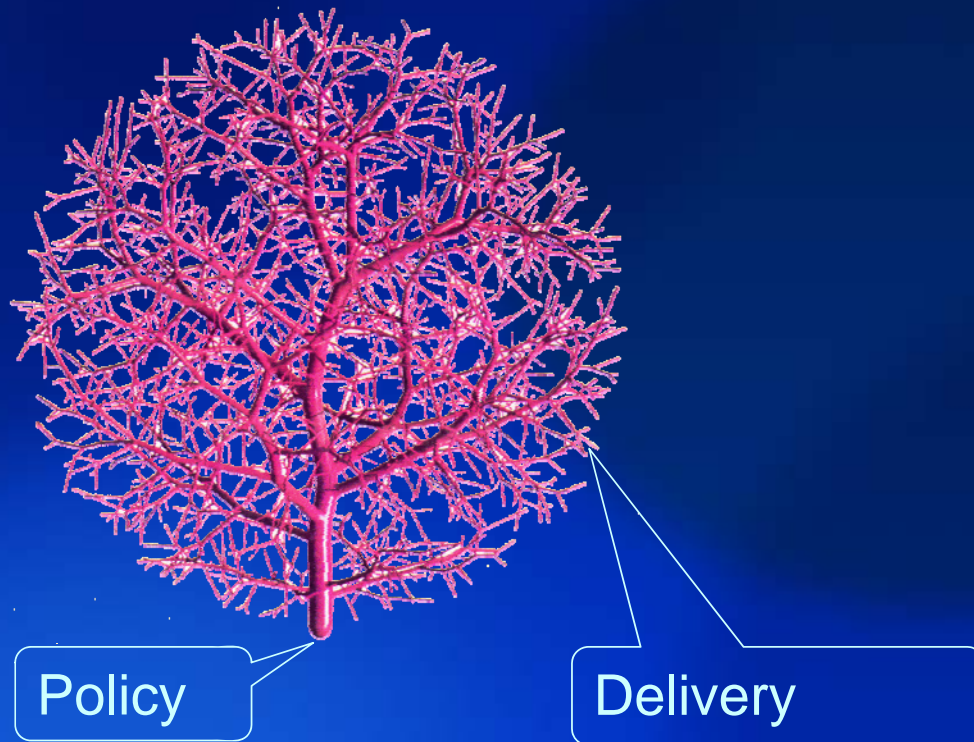




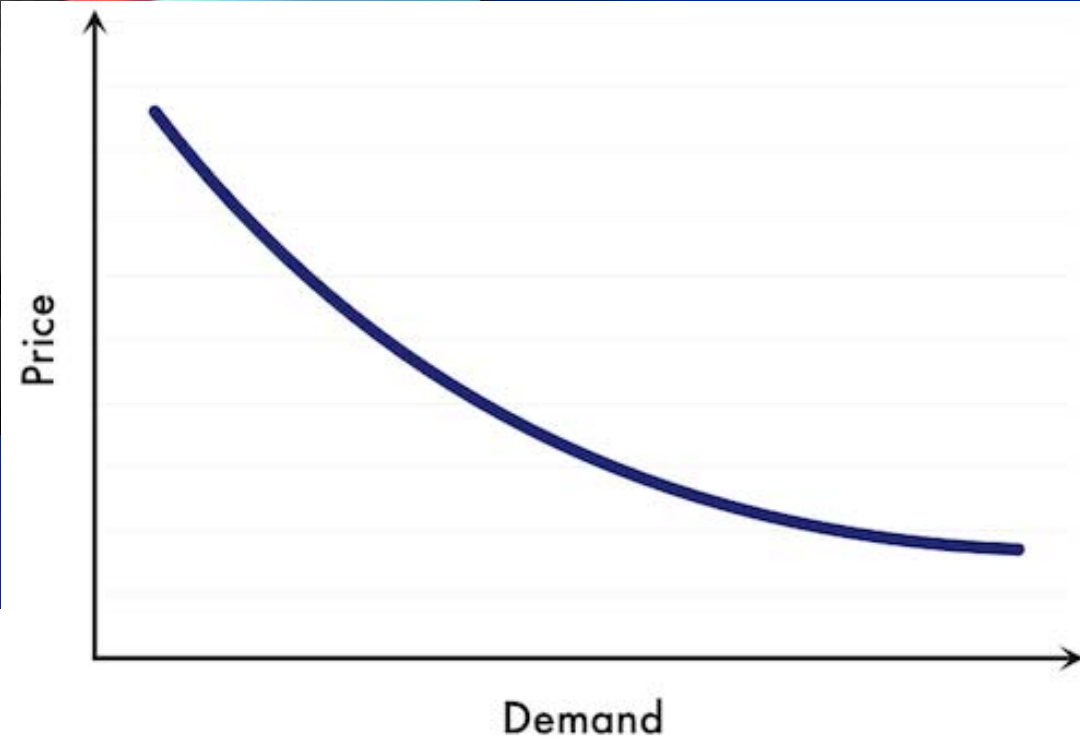
# 4 Evidence based delivery and evaluation



# Evidence based *delivery*



What kind of  
data do we  
need?



@ngruen1

# Data on what *causes* what

“Our success at Amazon is a function of how many experiments we do per year, per month, per week, per day....”

Jeff Bezos

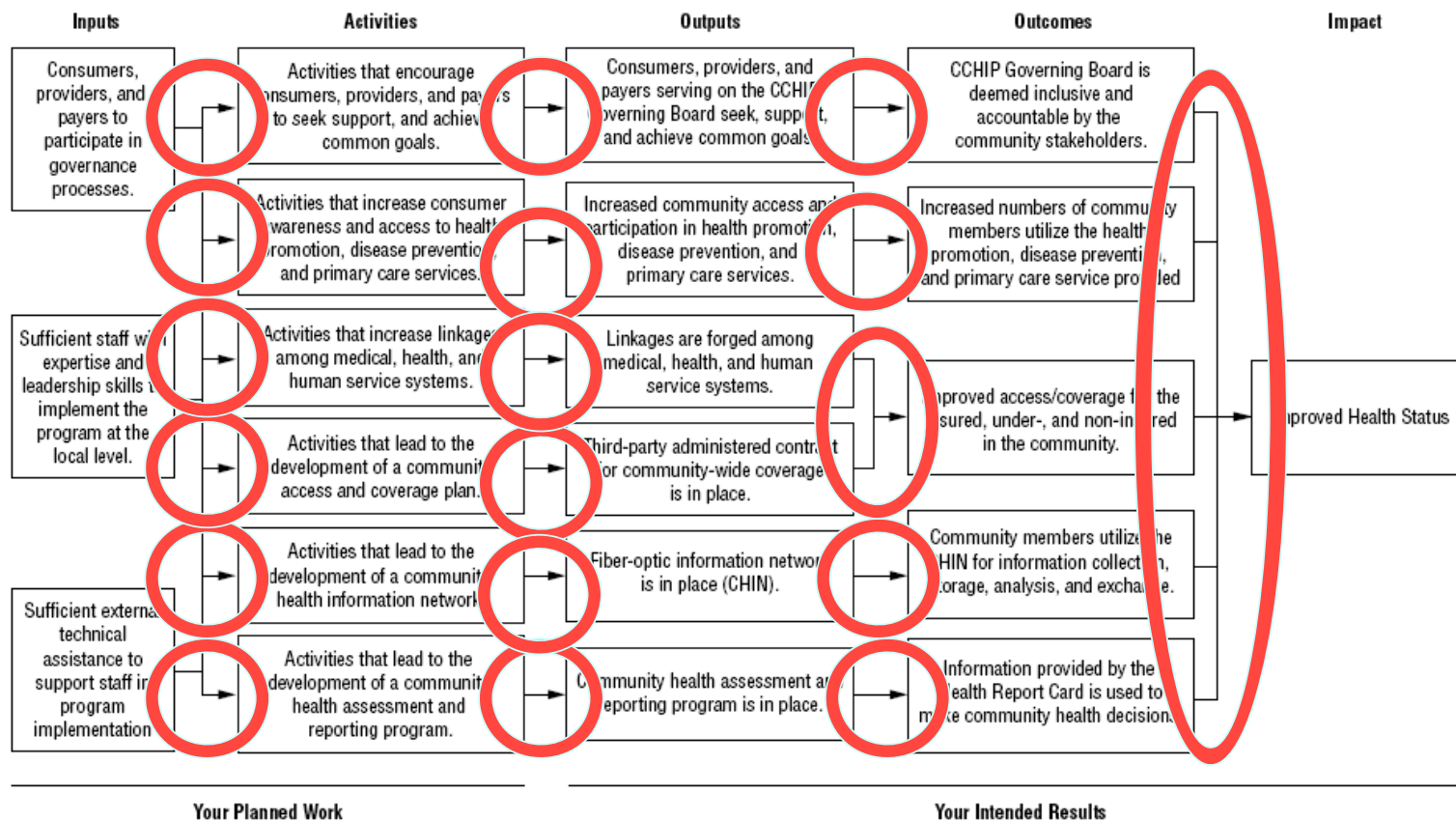
“Last year at Google the search team ran about 6,000 experiments and implemented around 500 improvements based on those experiments. The ad side of the business did about the same. Any time you use Google, you are in many treatment and control groups. The learning from those experiments is fed back into production and the system continuously improves.”

Hal Varian, chief economist at Google



# Program logic





Example of an Outcome Approach model (example drawn from the Calhoun County Health Improvement Program, funded under the Comprehensive Community Health Models of Michigan initiative).

ROUTLEDGE STUDIES IN GOVERNANCE AND PUBLIC POLICY

## The Politics of Evidence

From evidence-based policy to the good governance of evidence

Justin Parkhurst



One of the main conceptual holdovers from the world of evidence-based medicine has been the widespread, and often uncritical, embrace of so-called 'hierarchies' of evidence

Justin Parkhurst, *The Politics of Evidence: From evidence-based policy to the good governance of evidence* 2017



Lateral Economics

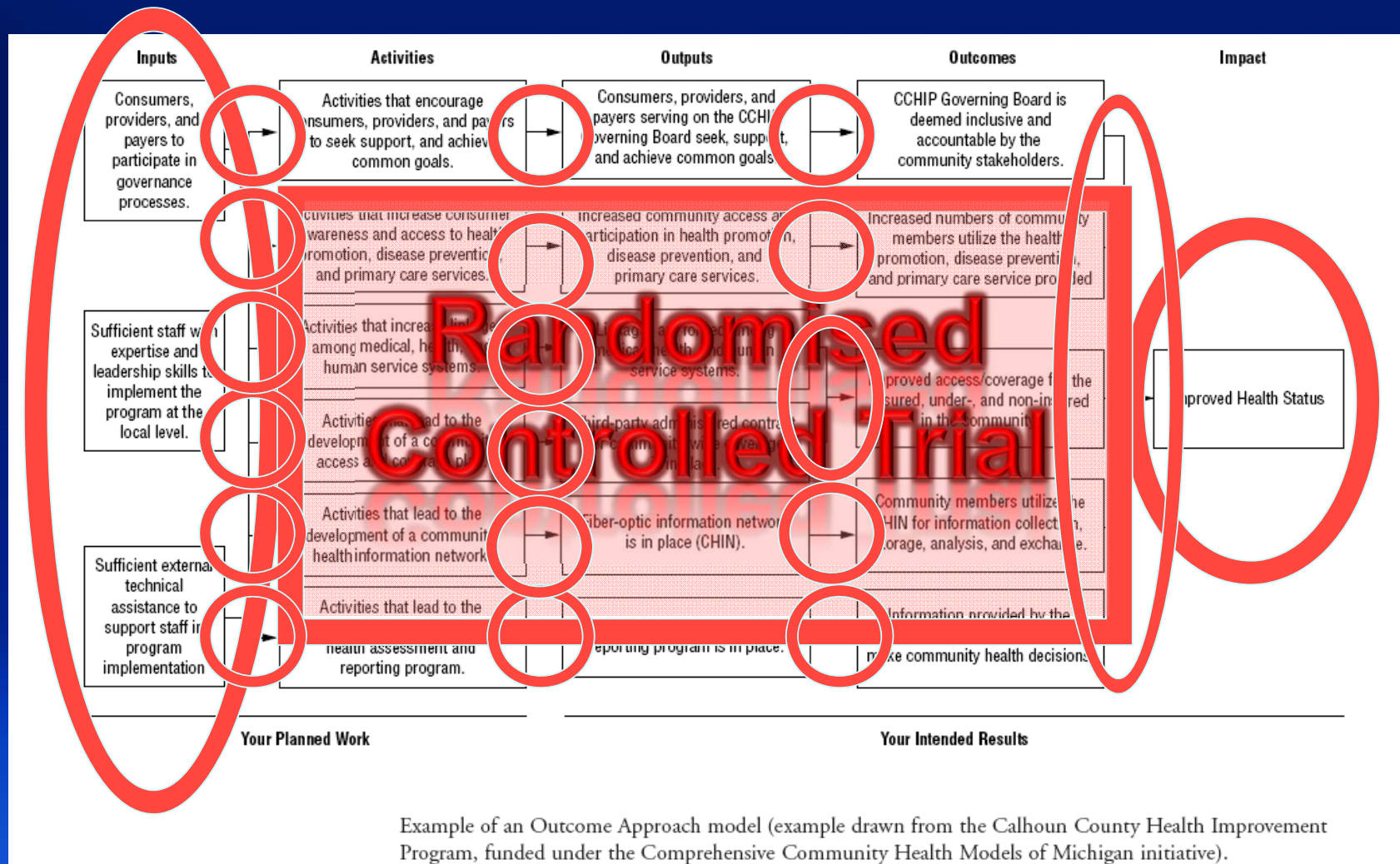
# RCTs can be important but they're thin



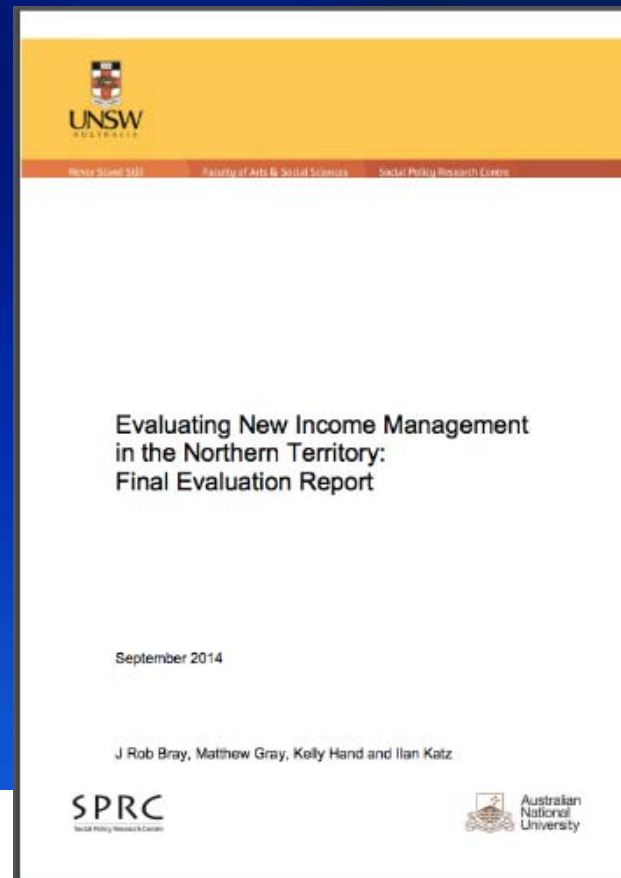
## Critics of overemphasis on RCTs

- Angus Deaton
- William Easterly
- Dani Rodrik
- Sanjay Reddy
- “Randomization is a metaphor and not a gold standard,” James Heckman
- And “Student’s” collaborator, the experimental maltster and barley farmer, Edwin S. Beaven.





# A small part of evidence based policy ...

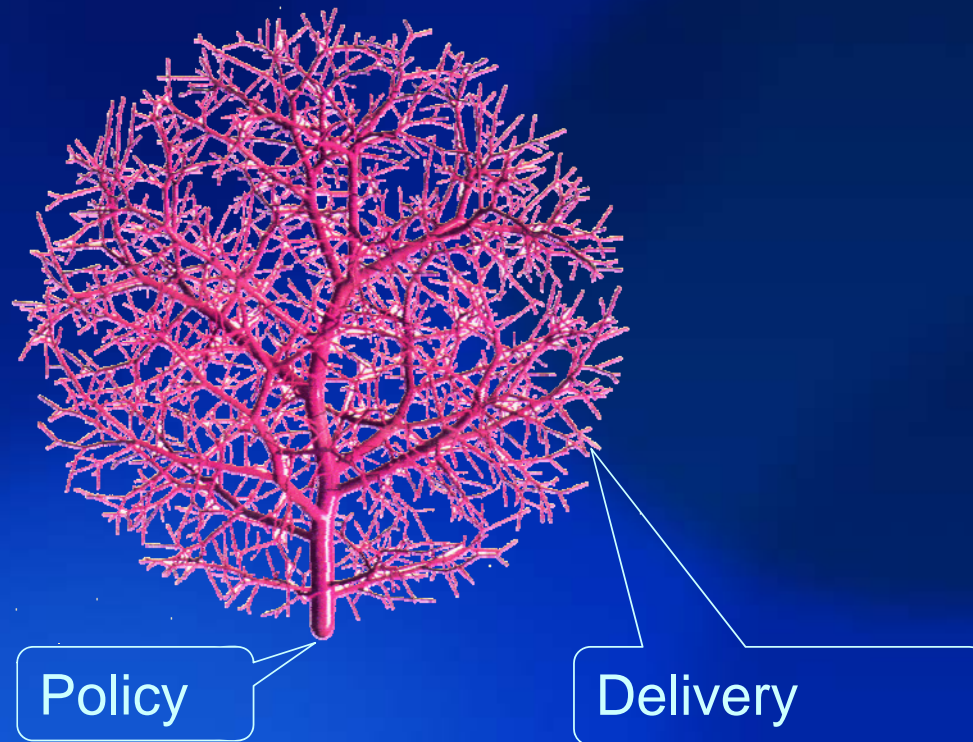


@ngruen1



Lateral Economics

# Evidence based *policy*



# Evidence based policy: What we're after

## Expertise

- To understand difficulties and maximise the chance to learn

Built round the practical needs of field workers to improve

## Causal data (A/B testing)

- To help us learn and improve

## Openness

- To build a community of practice and collaborative problem solving

## Independence

- To keep us honest, externally, internally, up, down

## Incentive compatibility

- To keep us trying

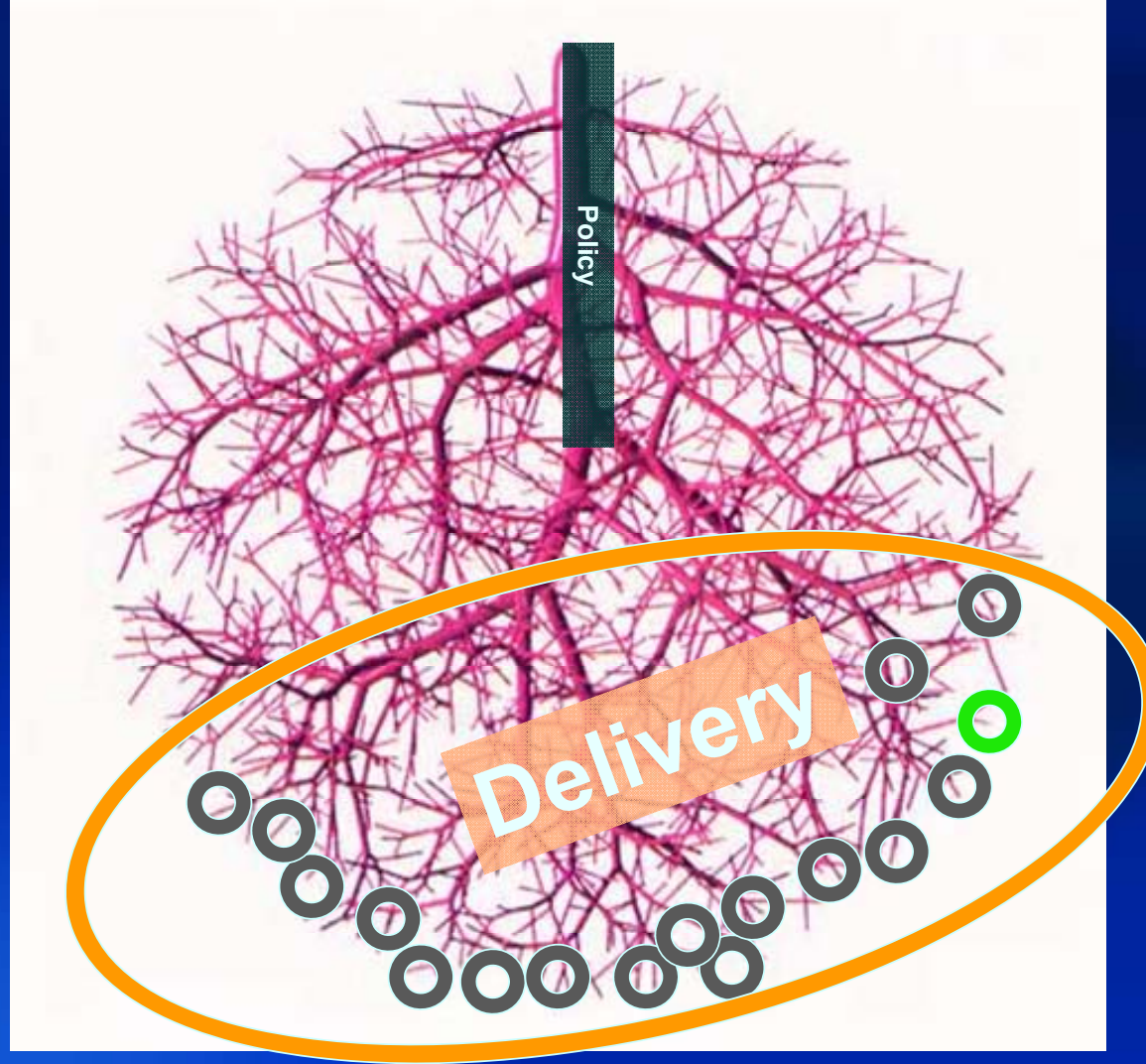
Evaluation planned and built in. Not retrospective and bolt on

- So it's efficacious





Learning goes upward =>



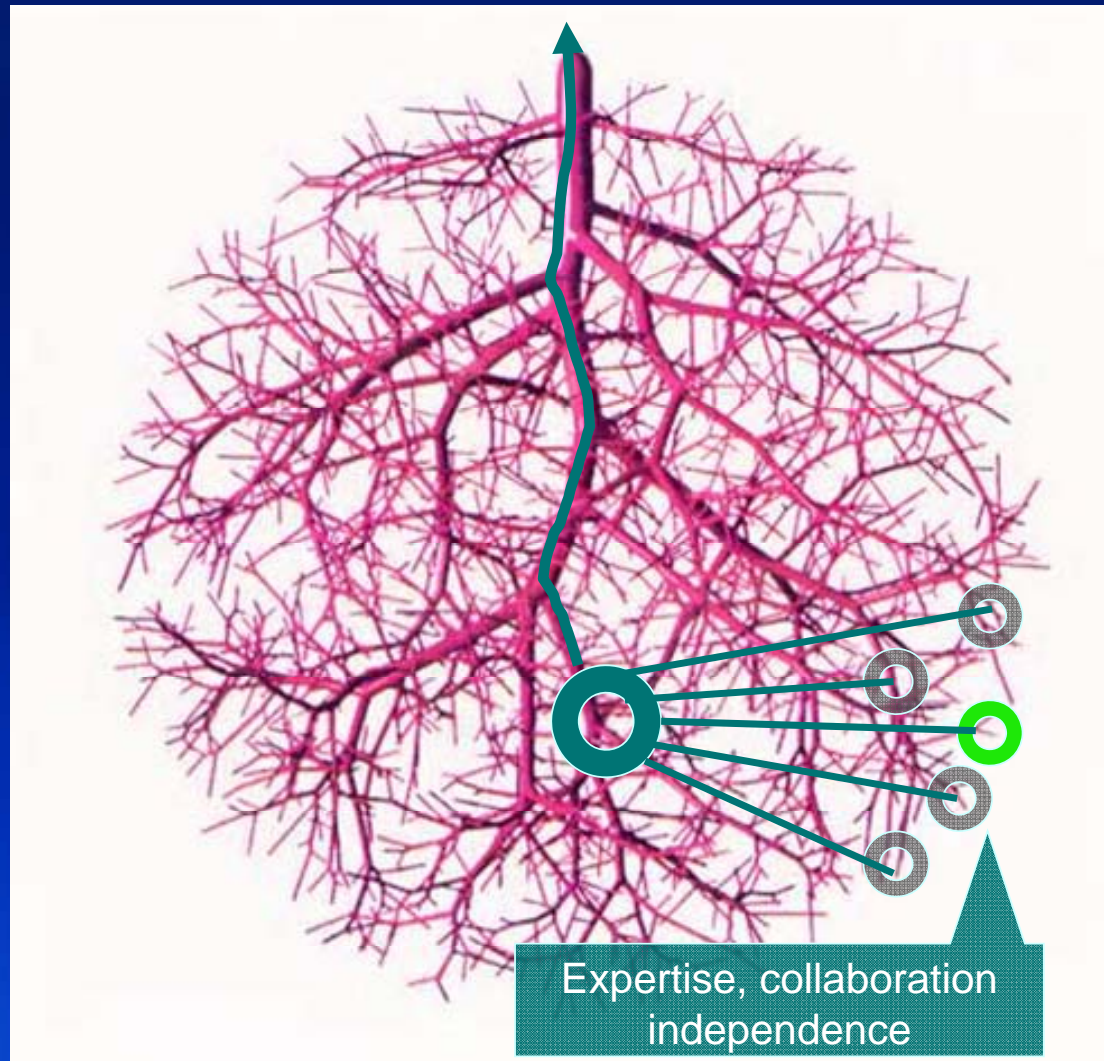
# Institutionalising evidence-based policy





Continual  
testing against  
the facts of the  
life world



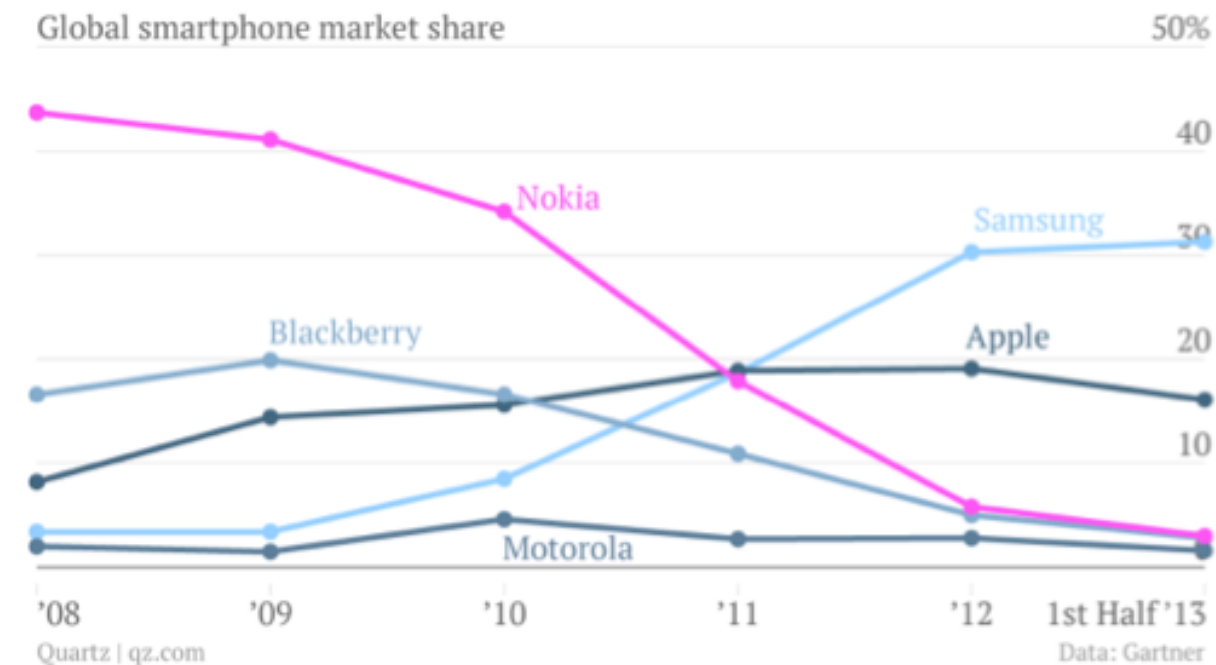




# The institutional imperative



**Figure 5.2: Markets adjust spontaneously to improvements in technology**

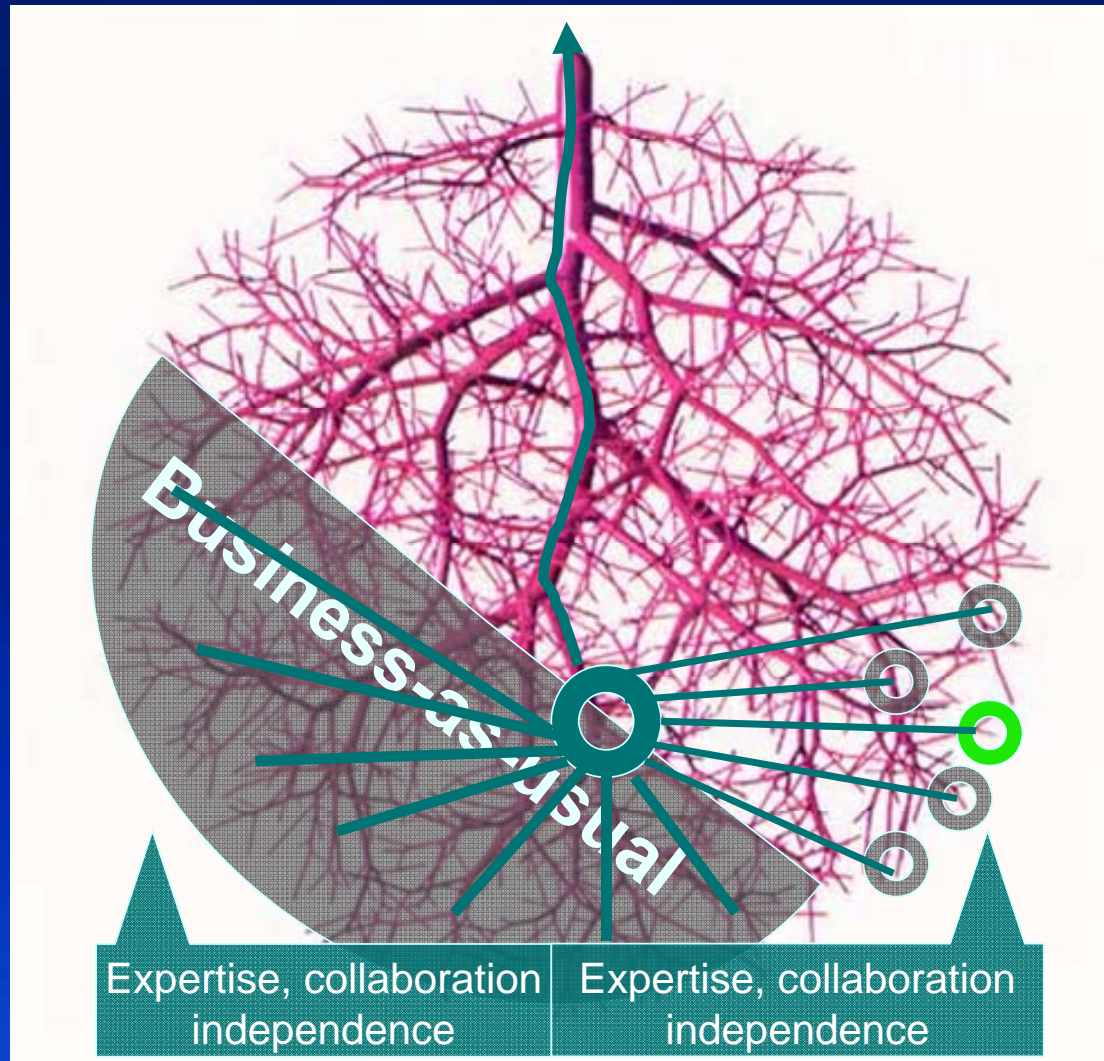


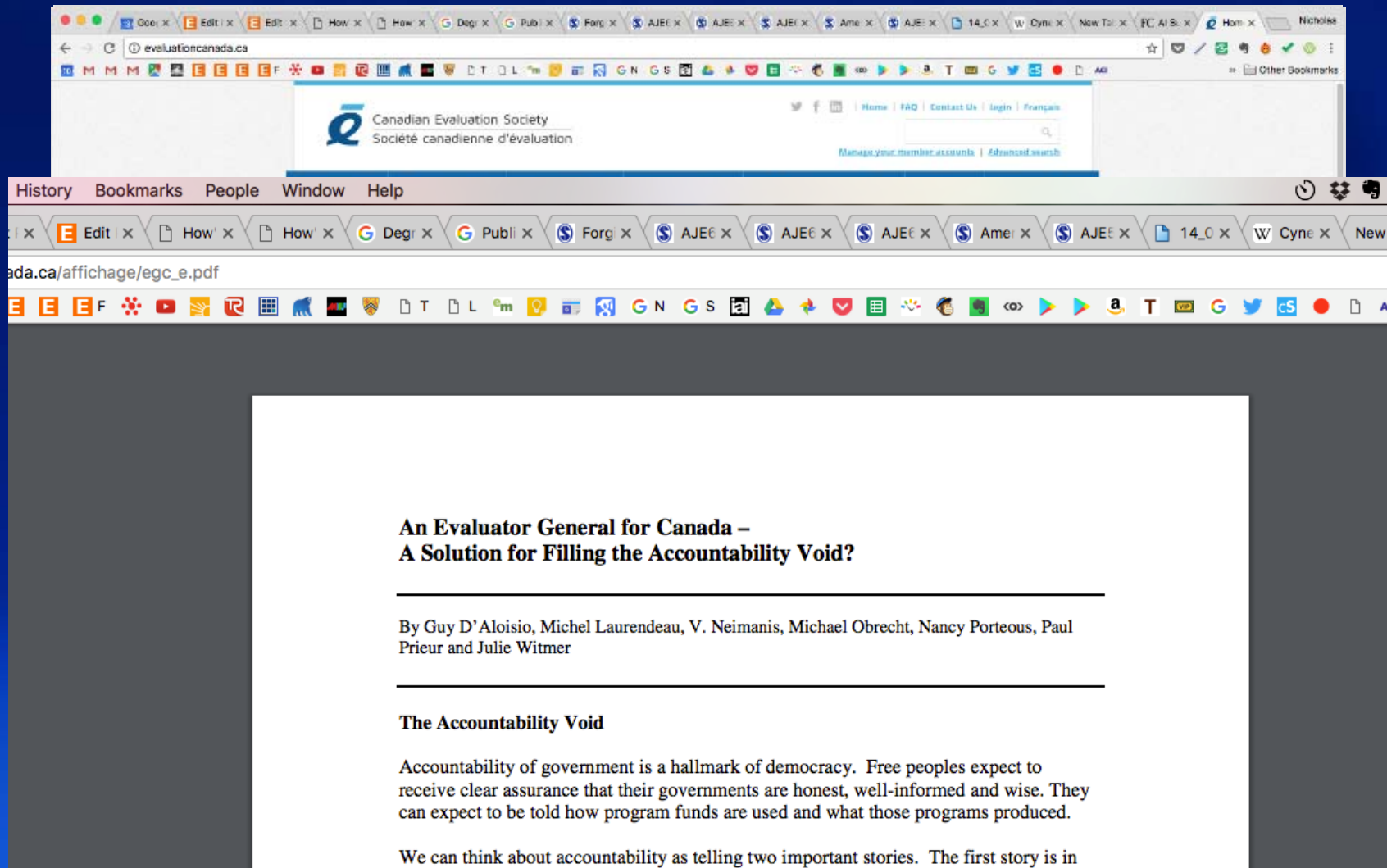
Source: <http://tinyurl.com/y825v9au>

Too much innovation remains at the margin of public administration. Opportunities are only half-seized; new modes of service delivery begin and end their working lives as 'demonstration projects' or 'pilots'; and creative solutions become progressively undermined by risk aversion and a plethora of bureaucratic guidelines.

Peter Shergold, 2013.



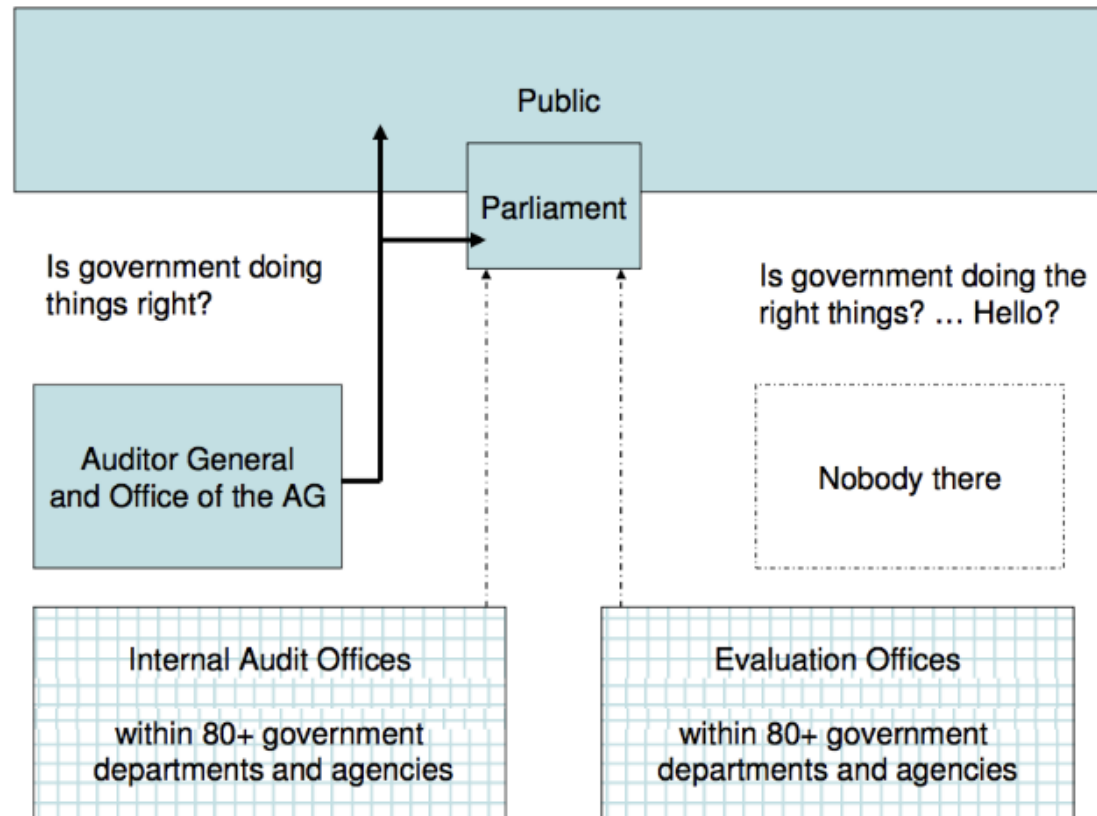


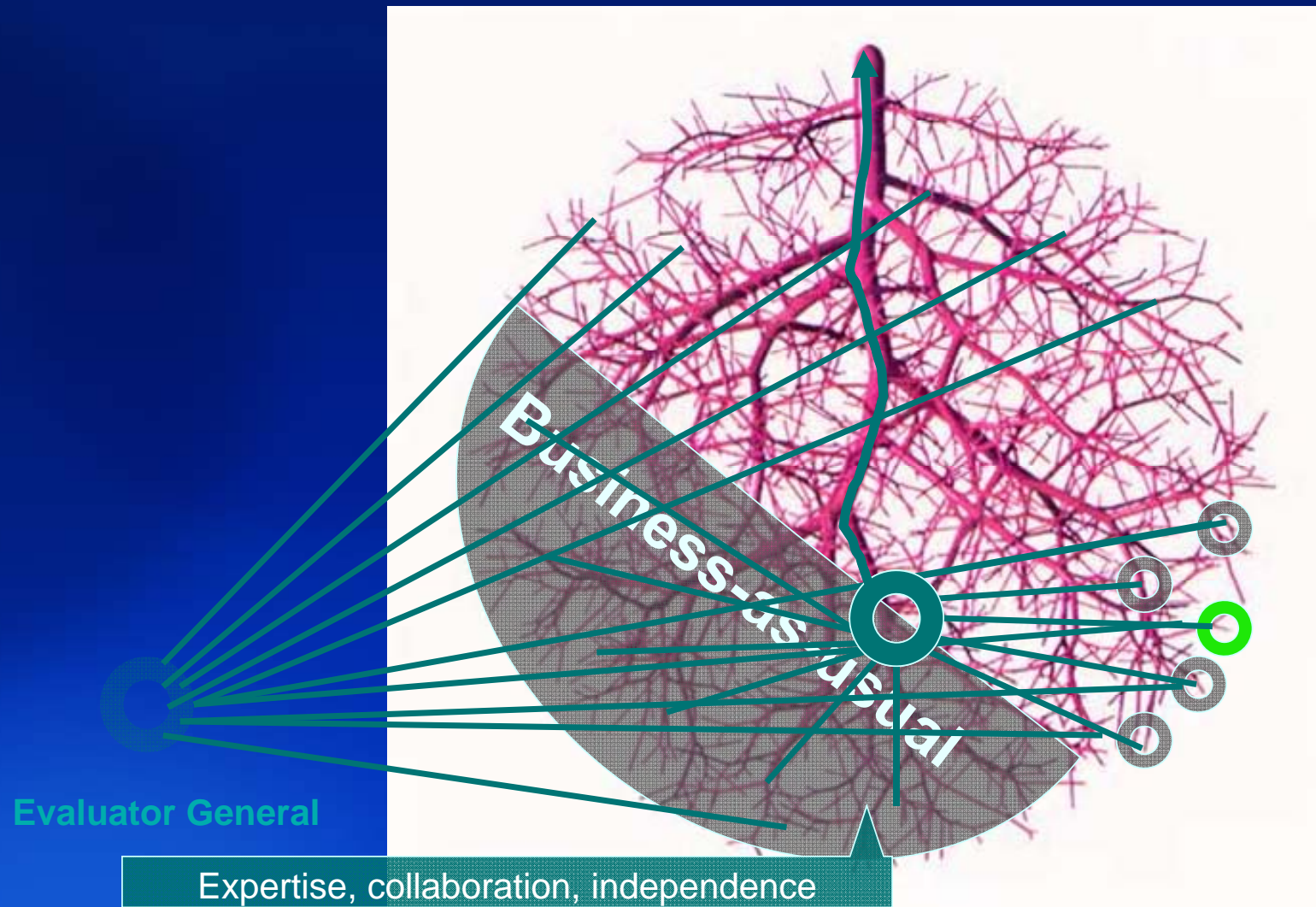


@ngruen1

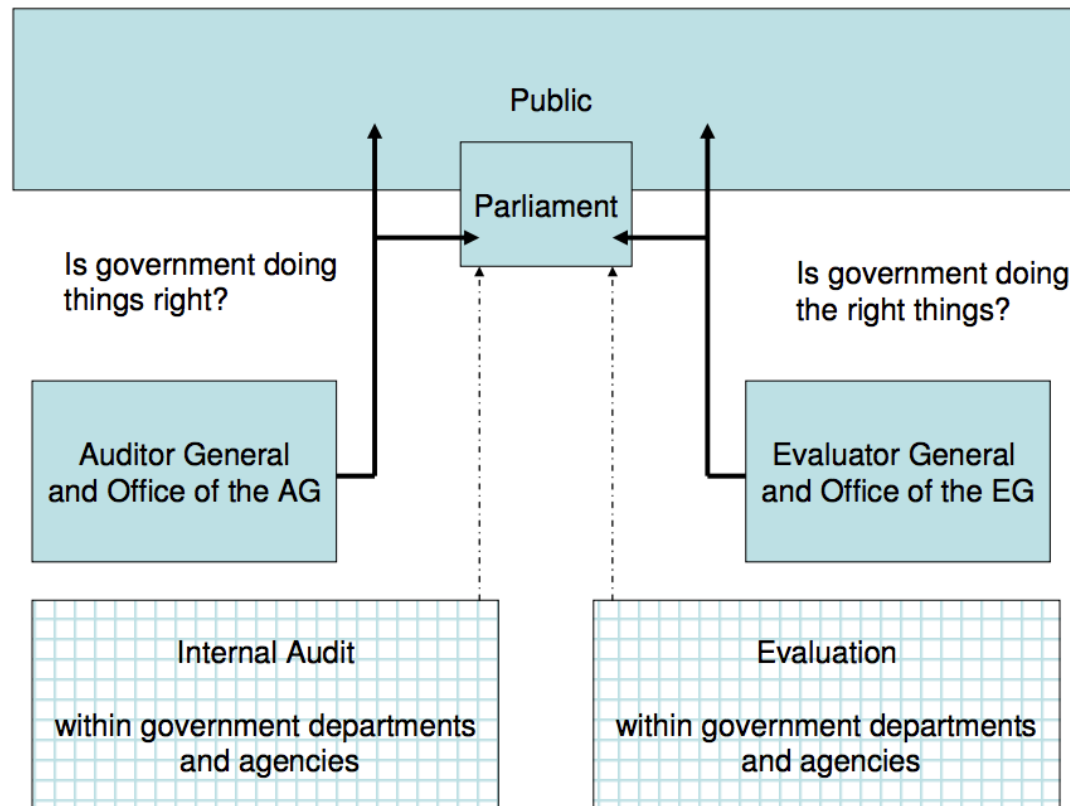


*Figure 1 - Current Situation: An Accountability Void*





*Figure 3 - Preferred Approach: An Evaluator General for Canada*



# Two systems

## Direct provision

- Integrity
  - Auditor Gen'l, Ombsmn
- Information
- Knowing what we're doing
- Understanding policy choices
  - PC
  - PBO

## Competitive provision

- Making choices
- Delivering services
  - Departments of state



# Accountability

It is not necessary to abandon the notion of being accountable for what has to be done but to return to the meaning and focus on systems of accountability that both justify [if it can be justified] and explain what has been done. This requires careful consideration of *who* is being held **accountable, to whom, for what, how, and with what consequences.**

More thoughtful and comprehensive approaches to accountability should demonstrably support good performance and encourage responsibility.

Patricia Rogers, ANZSOG



# Evidence based delivery: Good things to have

## Expertise

- To understand difficulties and maximise the chance to learn

Built round the practical needs of field workers to improve

## Causal data (A/B testing)

- To help us learn and improve

## Openness

- To build a community of practice and collaborative problem solving

## Independence

- To keep us honest, externally, internally, up, down

## Incentive compatibility

- To keep us trying

Evaluation planned and built in. Not retrospective and bolt on

- So it's efficacious



# Starting out

Grow expertise and independence of evaluation

- Can start very small
  - within government agencies

Identify some priority sectors (and/or regions)

- Indigenous policy
- Child protection

Set system targets with Auditor General reporting on them



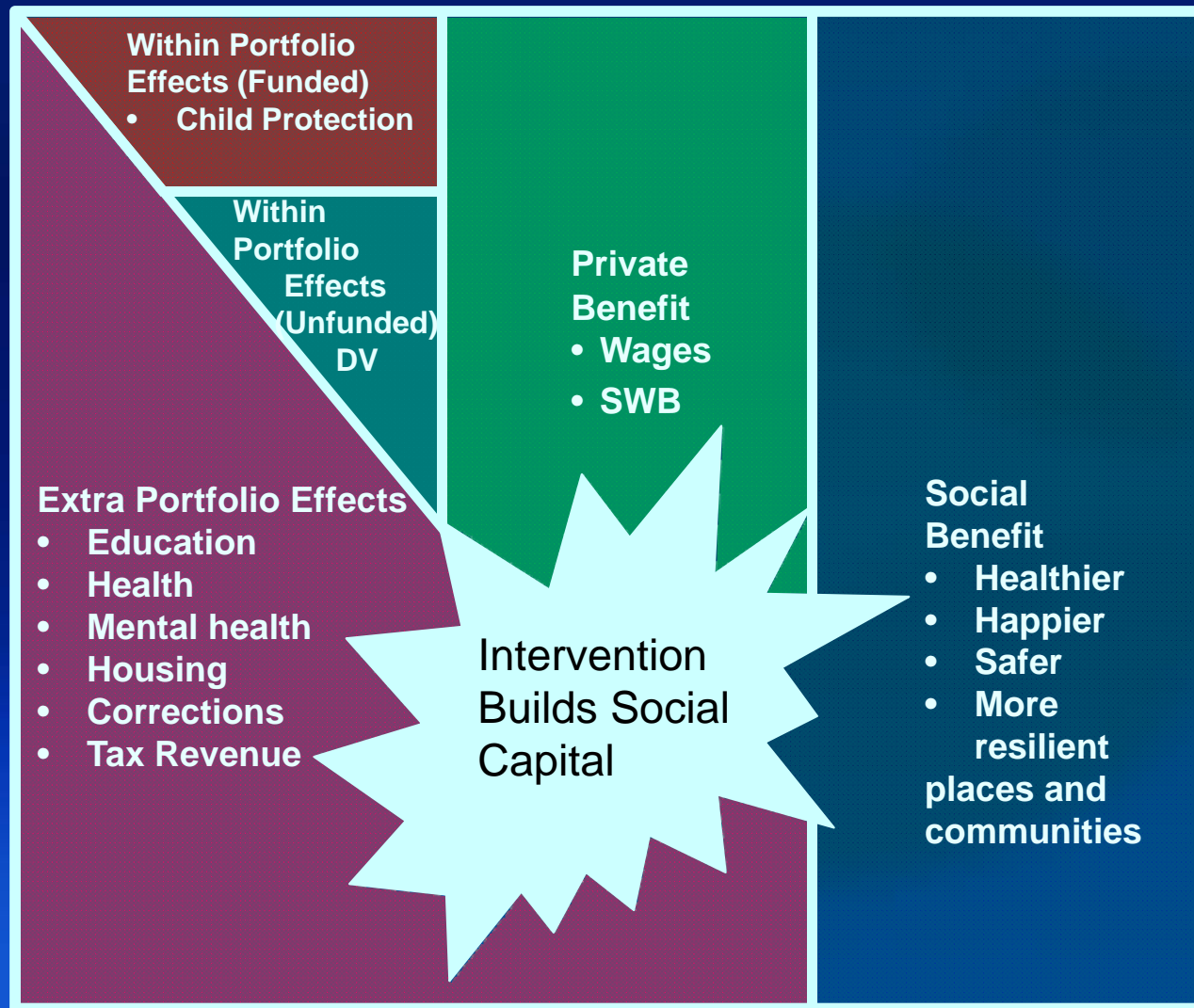
E ngruen@lateraleconomics.com.au

The end

@ngruen1



Lateral Economics







@ngruen1



Lateral Economics