

The National Cancer Informatics Program (NCIP) Hub

A platform for collaboration and sharing of data,
tools, and standards amongst the cancer
research community

Ishwar Chandramouliswaran

September 29 2014

National Institutes of Health (NIH)

Turning Discovery into Health

- Agency of the United States Department of Health and Human Services
- Primary agency responsible for biomedical and health-related research

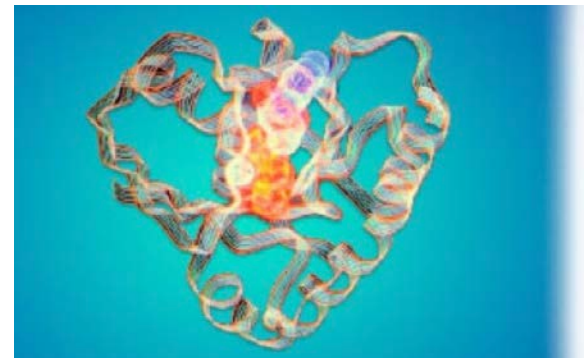
- Formed in 1887 - Hygienic Laboratory
- Headquarters in Bethesda, Maryland, U.S.
- Annual budget US \$30.9bn (as of 2010)
- Agency Executive is Dr. Francis Collins



- 27 Institutes/Centers, each focusing on particular diseases/body systems

National Cancer Institute

- Federal Government's principal agency for cancer
- Established under the National Cancer Institute Act of 1937
- Coordinates the 'National Cancer Program':
 - Research
 - Training
 - Education



1-800-4-CANCER

National Cancer Institute

Center for Biomedical Informatics & Information Technology

- Situated in the Office of the NCI Director
- Assess and meet informatics needs of NCI Programs
- Focus is on innovative programs and technologies supporting **open science** to accelerate research in the cancer informatics community



National Cancer Institute

Empowering Cancer Research via Science as a Service (SaaS)

GitHub
Channel

NCIP Open Source



at NCI

Supporting
Cancer Research
with Open Source

Visit ncip.nci.nih.gov
Wiki Pages

Open Development
Initiative

Democratize
Access

Cancer Cloud



NCIP Hub

Foster
Communities



National Cancer Institute

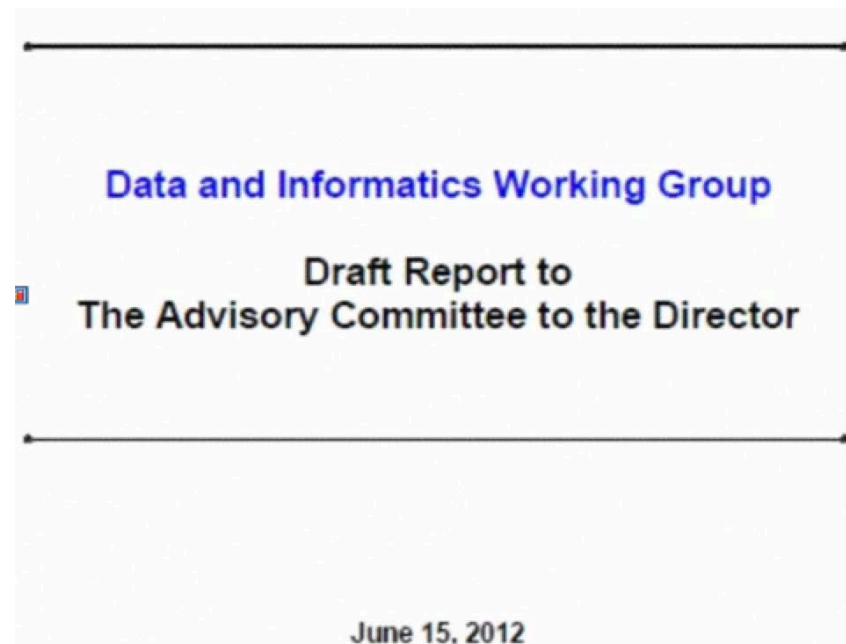
Realize Open Science to Accelerate Research

- Foster a **sustainable community** of expert members to support open innovation of tools, data and standards supporting cancer research.
- Promote mechanisms to **democratize access** to data, tools, and standards within NCI and across the cancer research community.
- Enable the scientific goals of the NCI through the **collaborative development** and support of informatics programs.

Align with NIH Informatics Objectives

Mission and vision of the Data Informatics Working Group

<http://acd.od.nih.gov/diwg.htm>



NIH Informatics Needs – Problems to Solve

- Locating the digital assets – breaking the silos
- Access to the asset – lowering the barrier
- Organizing, managing and processing/sustainability
- Accelerating new methods for analyzing (data)
- Sharing/collaboration - policies, practices, etiquettes
- Training researchers to use these assets effectively

Create an Adaptive & Collaborative Environment

Recommendations of the DIWG Committee:

- Promote data sharing through catalogs
- Support informatics methods & applications
- Build capacity via training in quantitative sciences
- Develop NIH-wide strategic plans
- Commitment to support these recommendations

The Big Data to Knowledge Program (BD2K)

Enable the biomedical enterprise to maximize the value of biomedical data



The screenshot shows the NIH Big Data to Knowledge (BD2K) website homepage. The header features the NIH logo and the text "NIH Big Data to Knowledge (BD2K)" with the tagline "Advancing Health and Discovery through Big Data". Below the header is a navigation menu with links for "FUNDING OPPORTUNITIES & NOTICES", "WORKSHOPS", "NEWS", "ABOUT BD2K", and "FAQs". The main content area is divided into several sections:

- DATA SHARING:** A central section with a blue background and a network diagram of people. It includes the text "The NIH Big Data to Knowledge (BD2K) new video on" and a "WATCH NOW" button.
- ADDS:** A section for the Associate Director for Data Science (ADDS), featuring a photo of Philip E. Bourne and a brief description of his role.
- PHI Bourne Blogs:** A link to the Associate Director's blog.
- NEWS HIGHLIGHT:** A section with two news items:
 - "Important Information for BD2K Centers RFA HG-13-999 Applicants" dated April 15, 2014.
 - "NIH Names Dr. Philip E. Bourne First Associate Director for Data Science" dated December 9, 2013.
- FUNDING OPPORTUNITIES & NOTICES:** A section with three funding opportunities:
 - "Input on Information Resources for Data-Related Standards Widely Used in Biomedical Science(RF) Response Date: September 30, 2014"
 - "Predoctoral Training in Biomedical Big Data Science (T32) April 2014"
 - "Revisions to Add Biomedical Big Data Training to Active Institutional Training Grants (T32) April 2014"
- CONNECT WITH US:** A section with links to "Join the BD2K Listserv" and "Follow us on Twitter".

<http://bd2k.nih.gov>

National Cancer Institute

NCIP Hub

A Collaborative for Informatics in Cancer Research

The screenshot shows the NCIP Hub website homepage. At the top left is the logo for NCIP HUB, with the text "A COLLABORATIVE FOR INFORMATICS IN CANCER RESEARCH" to its right. On the top right, there are buttons for "Login/Register" and "Get Help". Below the header is a navigation bar with "DISCOVER", "COMMUNITY", "ABOUT", and "SUPPORT" menus, and a search bar on the right. The main content area features a large graphic on the left showing stylized human figures on a grid, with the text "Join an NCIP HUB Group" and "Create your own group or join one of the communities". To the right of this graphic is a "Welcome to NCIP Hub" section with a red "Get Started" button. Below the main content are sections for "Twitter @NCI_NCIP" and "Events".

Twitter @NCI_NCIP

The screenshot shows two tweets from the NCIP Hub. The first tweet is from HUBzero (@hubzeroplatform) dated 24 Sep, with the text "Hot off the press: Nat'l Cancer Informatics Program hub building a community to accelerate cancer research, discovery bit.ly/1v0t9uA" and "Retweeted by NCIP". The second tweet is from NCIP (@NCI_NCIP) dated 25 Sep, with the text "Registration is open for Rat #Genomics and Models Conference, December 1-4, @WellcomeTrust Genome Campus bit.ly/1vklvlv".

Events

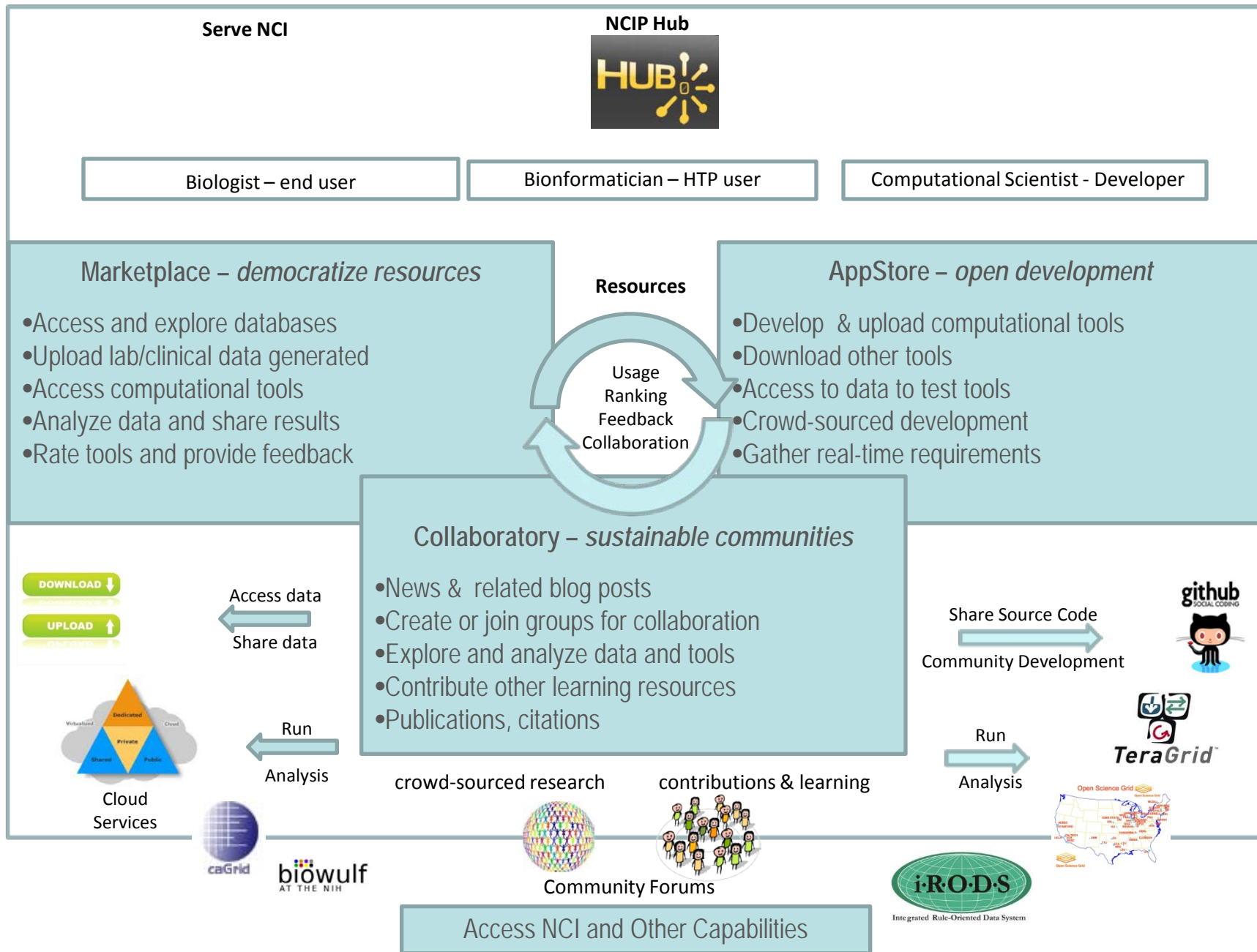
- 28** SEP **HUBhub Conference 2014**
- 02** OCT **3rd Annual Biomedical Informatics Symposium**
- 08** OCT **Beyond the Genome: Cancer Genomics**
- 28** OCT **Introduction to Next Generation Sequencing - Training**

NCIP Hub, A 360° Approach to Developing and Sustaining NCI Informatics Resources

- Serve as front-end to the cancer informatics assets
 - Access cancer related data/resources
 - Support open development projects
 - Marketplace for relevant analysis tools
- Enable sustainable communities
- Serve as a model for sharing and as an educational resource
- Enhance transparency and incentivize collaboration

VISION: NCIP Hub

Serve larger community



Why HUBzero?

- ✓ Broad set of features for scientific community
- ✓ Worked for diverse scientific domains
- ✓ Serves both individual and power users
- ✓ Provides transparent analytics
- ✓ Open source; NSF funded, sustainable research project

- ✓ Aligns with program & cancer community needs
 - Establish catalog of informatics resources
 - Content driven by the community
 - Support training needs
 - Reduce redundancies & enhance transparency
 - Reduce costs
 - Model for collaboration (internal & external)
 - Opportunities to innovate

Where are we today?

- Community site setup www.nciphub.org
- Hosted at Purdue University
- Understanding users & use at other federal agencies
- User group & webinars to teach & learn
- Lead user activities (test drives) underway
 - 500 members, 40 groups
 - 130 resources

Table 3: Miscellaneous Statistics

Domains Served	1,140
Cummulative Interactive User Sessions ¹⁰	4,811
Cummulative Session Time	635 days
Visitors ¹¹	10,506
Visits ¹²	38,230
New Accounts	409
Max User Logins	41 users on 2014-03-20
Web Server Hits	3,047,175

Quantitative Imaging Network: Collaborative development and algorithm evaluation

- Share (open source) tools for image analysis
- Develop & deploy statistical analysis & metrology tools
- Co-locate data & tools
- Evaluate the performance of quantitative algorithms
- Perform algorithm comparisons
- Encourage network member participation
- Make recommendations for best practices

Informatics Technology for Cancer Research: Communication & Collaboration

- Serve as platform for **communication** within ITCR
- Establish a **central access point** program generated data, tools & other resources
- Provide **individual project management** support
- Promote **inter-project collaboration**

Data (meta) Standards

Create Awareness, Discuss Gaps & Promote Adoption

STANDARDS PORTAL (PILOT)

The Standards Portal is envisioned as a potential solution for bringing together standards of interest in a unified and meaningful way.

TCGA CDE STANDARDS

TCGA version 2.6 CDEs are the standard elements in validation of clinical data, available for download at the Data Matrix. This resource presents the CDEs for search and exploration, with materials to clarify the CDEs and describe best practices for adopting them.



ADVERSE EVENTS STANDARDS

This updated version of naming and grading of adverse events in cancer clinical trials is the result of collaboration among academic, government and pharmaceutical industry clinical investigators. The effort is based on the need for reproducible, consistent and readily usable adverse events reports.



REQUEST AUTHORIZATION TO UPLOAD

Does your community want to upload preferred standards to the Standards Portal?



EXPLORE STANDARDS USING DATAVIEWS CDES, CRFS AND VOCABULARY

Use these DataViews to Explore the Standards Portal!

- [CDEs Complete View →](#)
- [CDEs Common Data Fields →](#)
- [CRFs Complete View →](#)
- [CRFs Common Data Fields →](#)
- [Preferred Content Vocabulary →](#)

Put DataViews on all your resource and group pages! Here is a list of all DataViews at the Standards Portal

[All Standards Portal DataViews →](#)

You can filter DataViews by Group:



DISCUSSION AROUND STANDARDS

This group site supports public discussion about discovery of standards, with gateways for easy access to all standards portal groups and resources.



The Standards Portal is a community-driven dynamic web site that centralizes access to valuable and practical information around standards to facilitate adoption and prioritization.

Our collaborative environment provides a means for research communities to identify and discuss gaps and overlaps for the evolution of preferred content, with powerful support for content sharing, search and discovery.

Nanotechnology Working Groups/Initiatives

- Nanotechnology working group meetings
 - Collaborative paper development
- National Nanotechnology Initiative: www.nano.gov
 - Discussion boards on data readiness levels
- Introduced the nanoHUB to the nanoEHS community
<https://us-eu.org>

Partnerships to Advance Cancer Health Equity

- Program in Center to Reduce Cancer Health Disparity
- All about building partnerships between minority institutions and cancer centers
- Collaboration for scientific discovery and academic learning

Establish and sustain a **CPACHE Community Forum** in NCIP Hub

The Oncology Models Forum: Open Interactive Platform for Translational Research

- Establish a knowledgebase supporting animal models of human cancer
- Provide an educational resource
- Enable
 - collaborations,
 - demonstrate best practices, and
 - share educational materials
 - better reproducibility of results

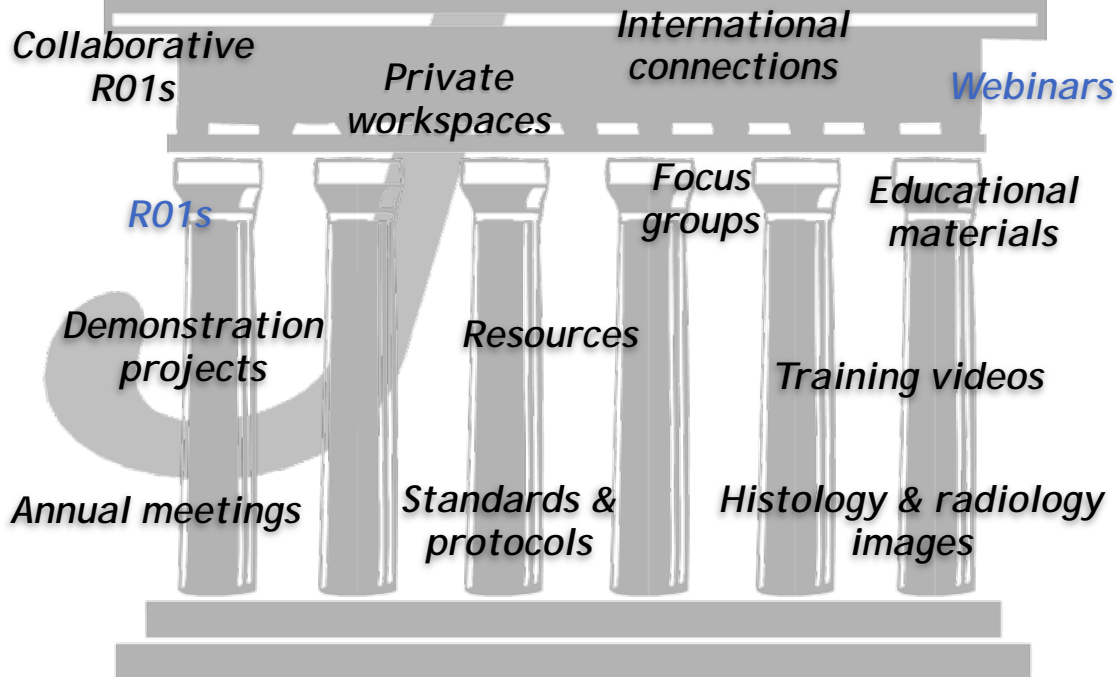


NCIP | HUB

Informatics Tools for
Cancer Research

Quantitative Imaging
Network

Oncology Models Forum



Challenges Ahead

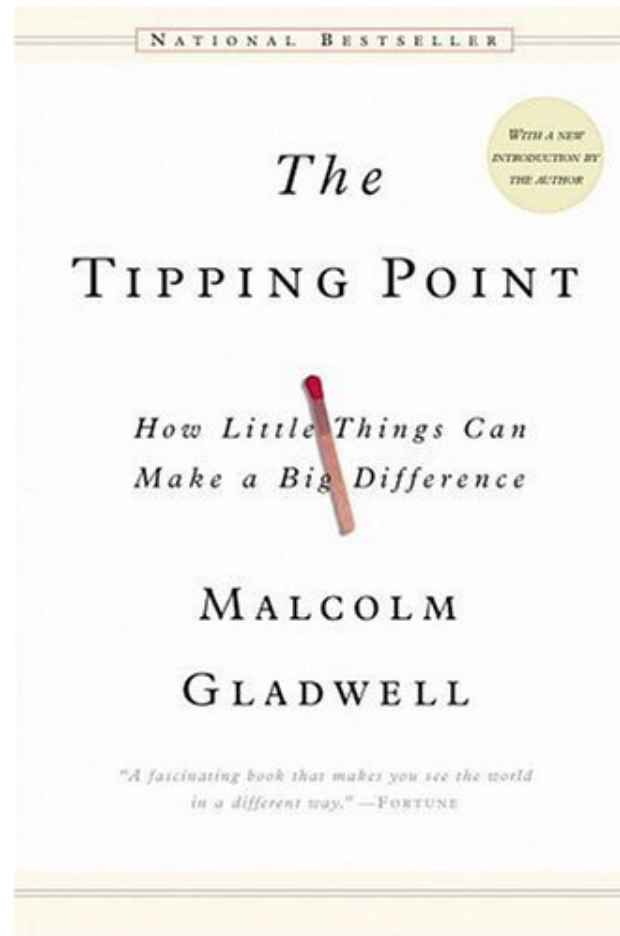
Community:

- Content to drive a vibrant community
- Identify the right partners/community champions
- Platform adoption, engage & incentivize users
- Encourage open-ness
- Co-exist with other platforms, tools

Operational:

- Acquire HUBzero admin & web development skills
- Comply with Federal Information System Requirements
- Lower the barrier for community utility

What is the NCIP Hub Tipping Point?



National Cancer Institute

Ishwar Chandramouliswaran

Email

: ishwar.chandramouliswaran@nih.gov

NCIP Hub

: <https://nciphub.org/members/1100>

LinkedIn

: www.linkedin.com/in/ishwarc/

Twitter

: <https://twitter.com/ishwartweets>

