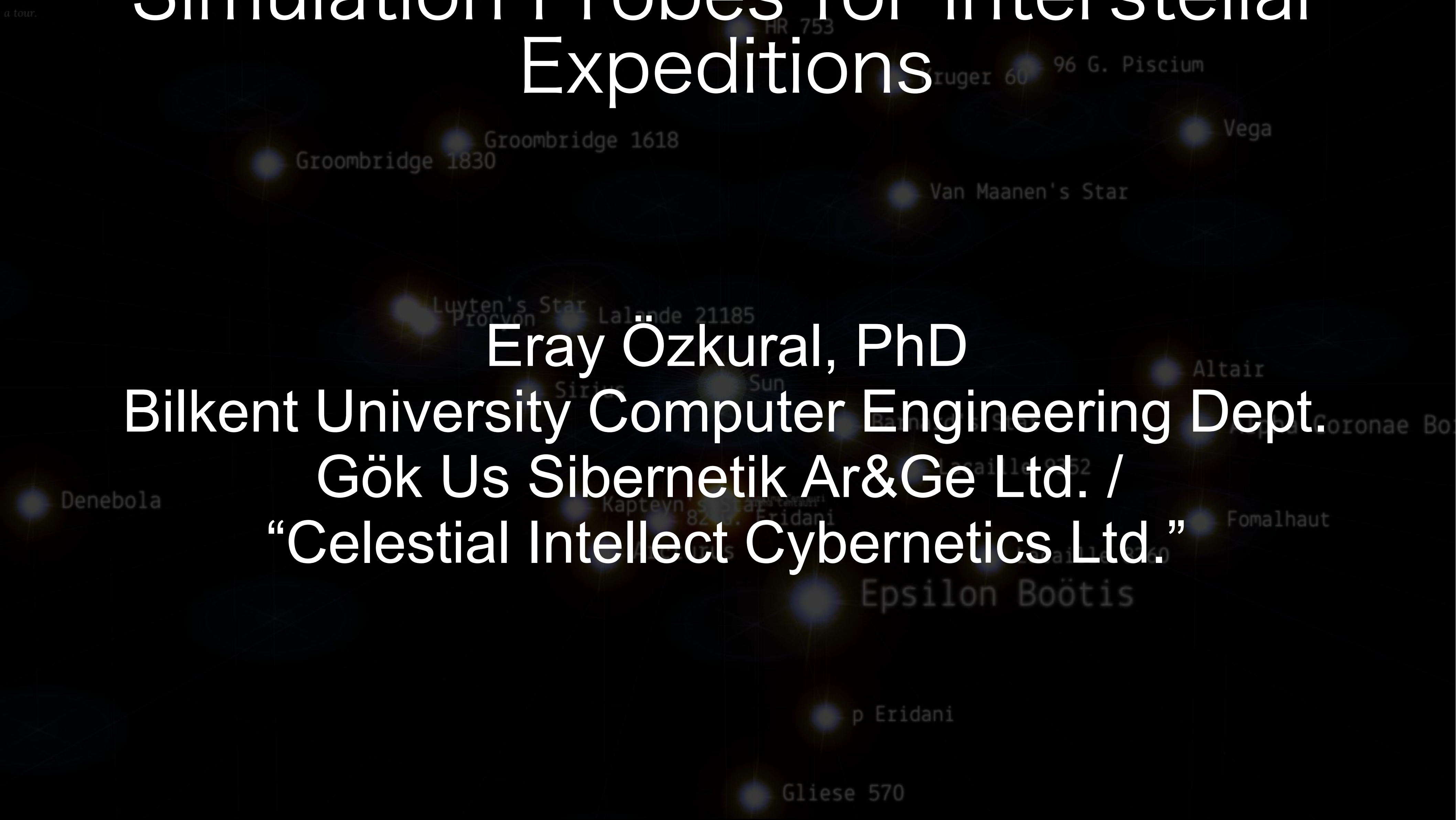
Artificial Intelligence and Brain Simulation Probes for Interstellar Expeditions

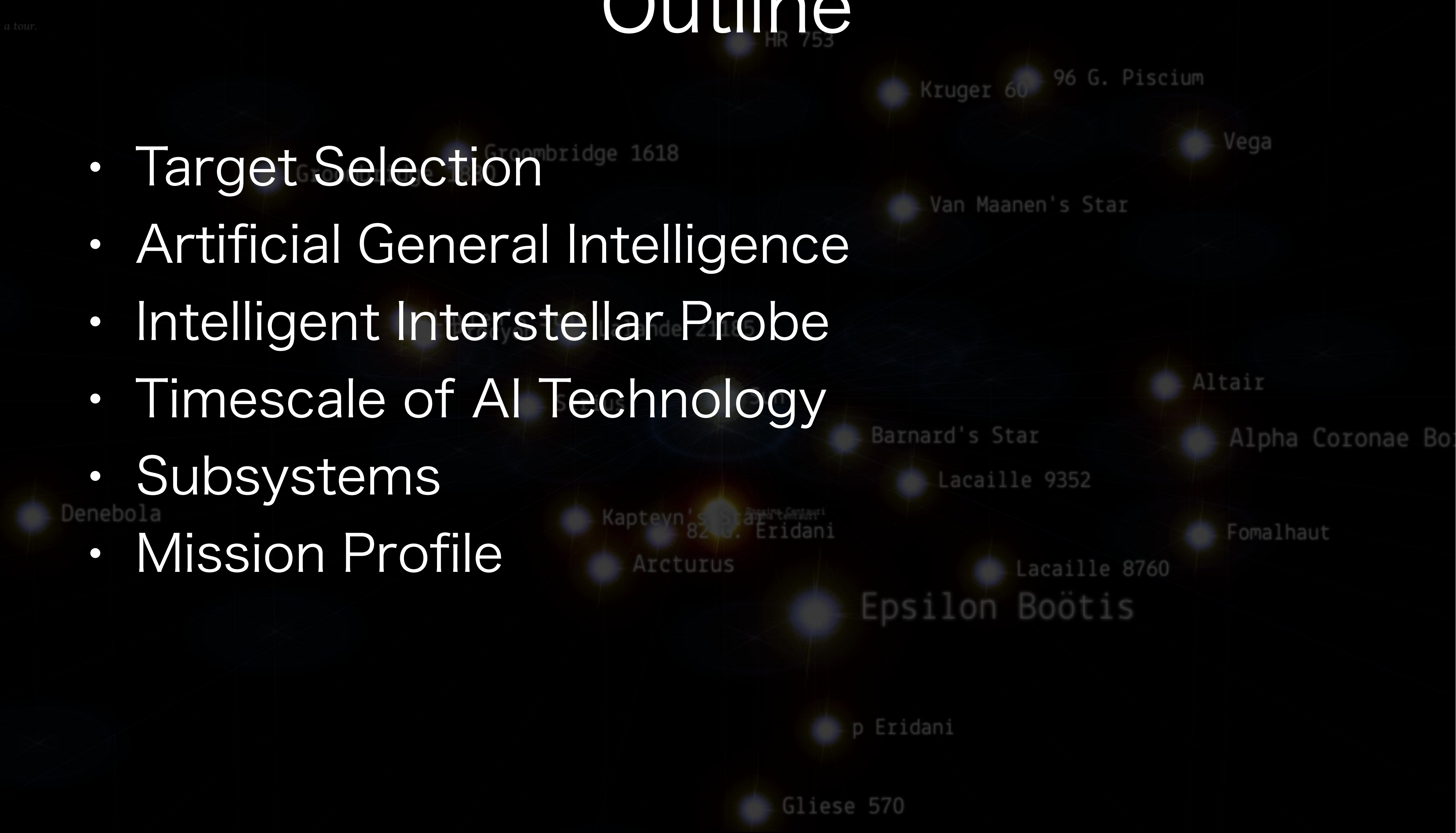


Eray Özkural, PhD
Bilkent University Computer Engineering Dept.
Gök Us Sibernetik Ar&Ge Ltd. /
“Celestial Intellect Cybernetics Ltd.”

100 YEAR STARSHIP™
sigma 2398 11.5 LY

Outline

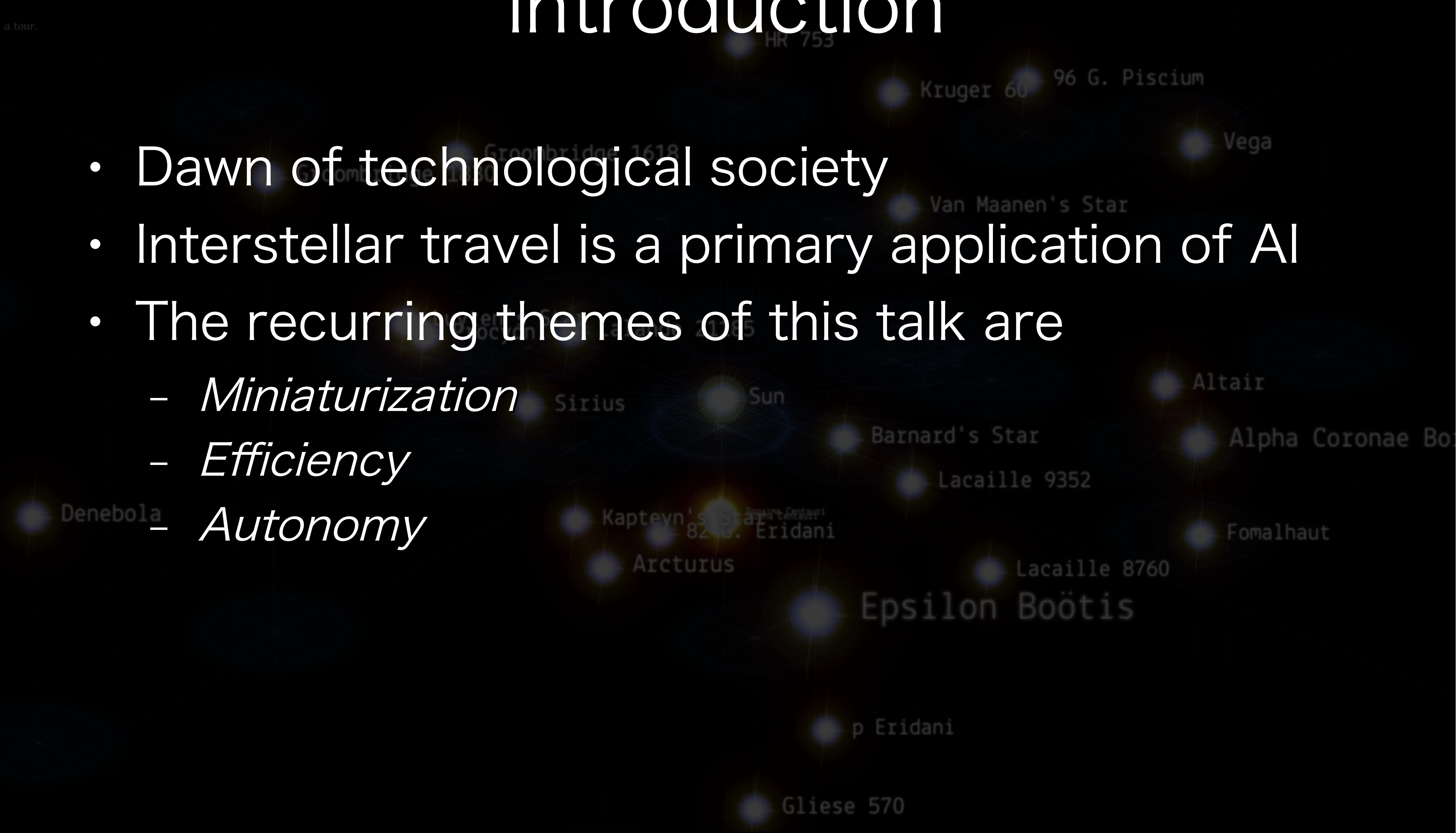
- Target Selection
- Artificial General Intelligence
- Intelligent Interstellar Probe
- Timescale of AI Technology
- Subsystems
- Mission Profile



100 YEAR STARSHIP™
sigma 2398 11.5 LY

Introduction

- Dawn of technological society
- Interstellar travel is a primary application of AI
- The recurring themes of this talk are
  - *Miniaturization*
  - *Efficiency*
  - *Autonomy*



Local Stellar Neighborhood

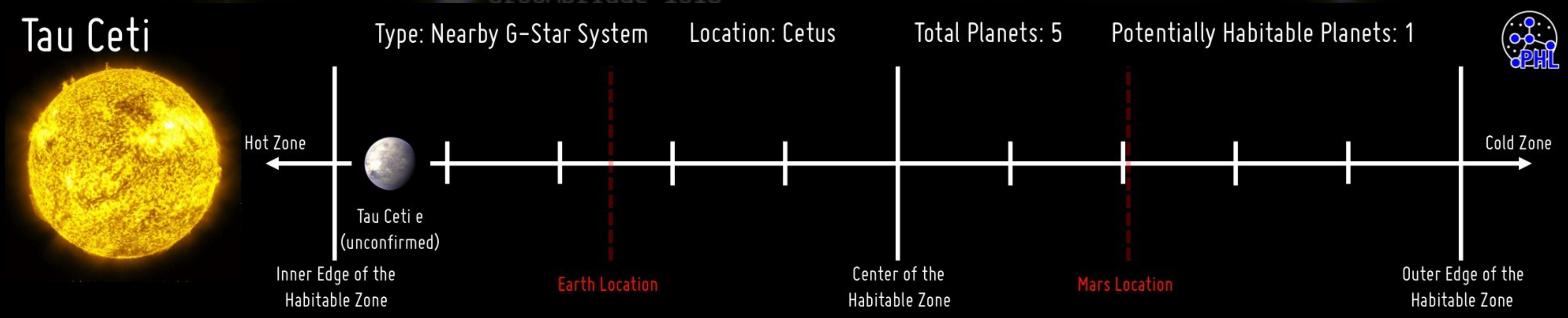
100 YEAR STARSHIP™
sigma 2398 11.5 LY

Targets: Nearby Systems with Habitable Zone Planets

Name	Distance (ly)	Planets	Planets in HZ	Max. ESI
Gliese 667C	22	6 (7)	3	0.82
Gliese 581	20.2	5	1	0.81
Tau Ceti	11.9	5	1 (2)	0.77
Gliese 163	48.8	3	1	0.74
HD 40307	41.7	3	1	0.72

- Gliese 667C: recently discovered super-Earths
  - Three planets
- Tau-Ceti:
  - Two possible planets

Target Selection



Tau Ceti

Type: Nearby G-Star System

Location: Cetus

Total Planets: 5

Potentially Habitable Planets: 1

Hot Zone

Tau Ceti e (unconfirmed)

Inner Edge of the Habitable Zone

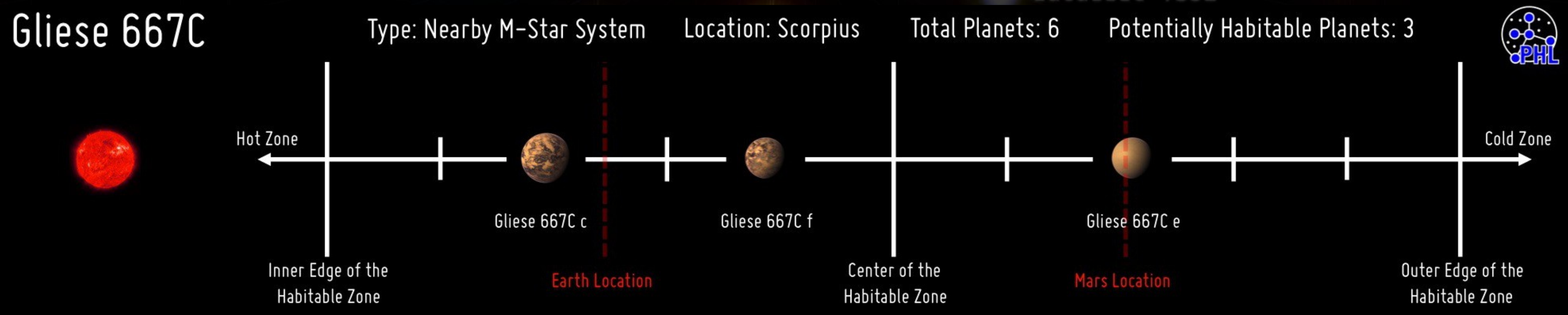
Earth Location

Center of the Habitable Zone

Mars Location

Outer Edge of the Habitable Zone

Cold Zone



Gliese 667C

Type: Nearby M-Star System

Location: Scorpius

Total Planets: 6

Potentially Habitable Planets: 3

Hot Zone

Gliese 667C c

Gliese 667C f

Gliese 667C e

Inner Edge of the Habitable Zone

Earth Location

Center of the Habitable Zone

Mars Location

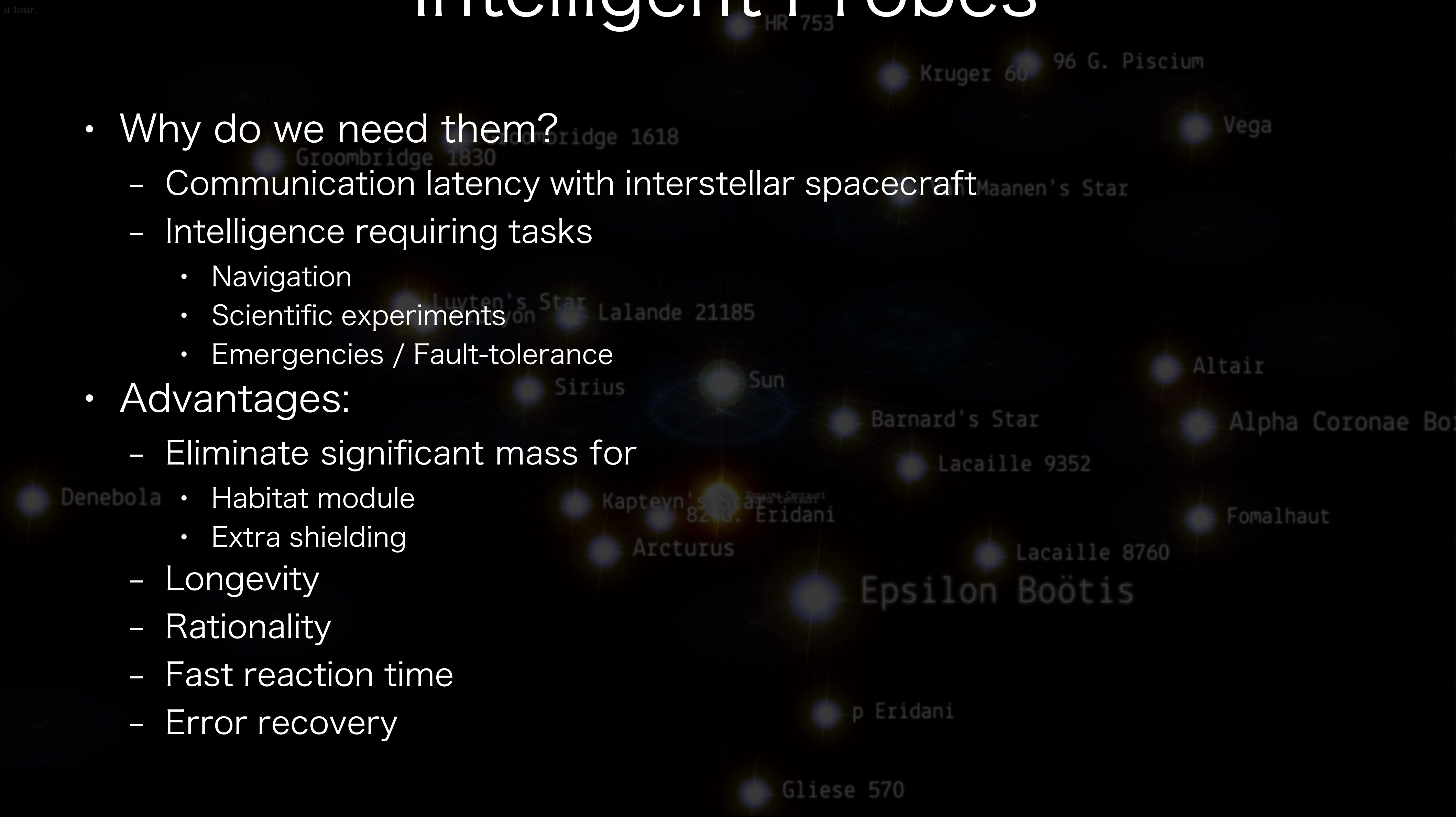
Outer Edge of the Habitable Zone

Cold Zone

100 YEAR STARSHIP™

sigma 2398 11.5 LY

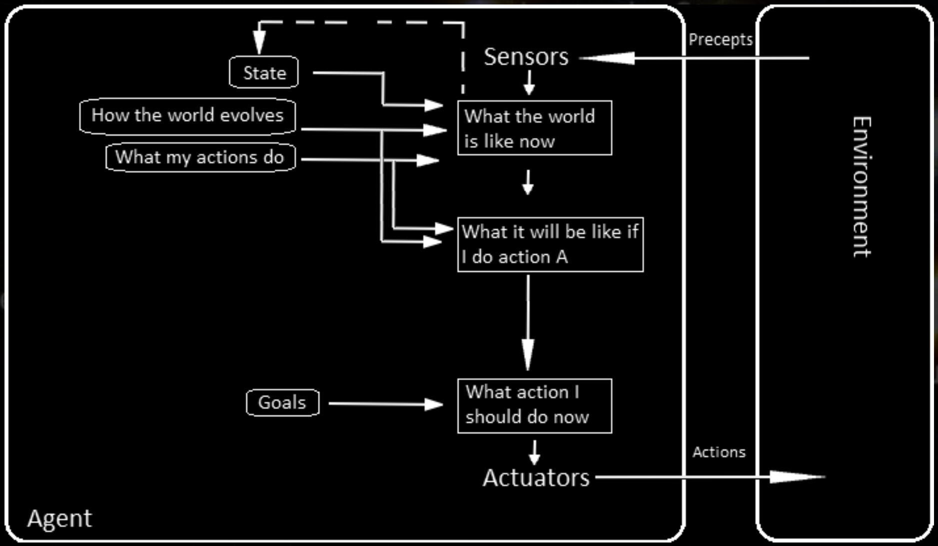
Intelligent Probes



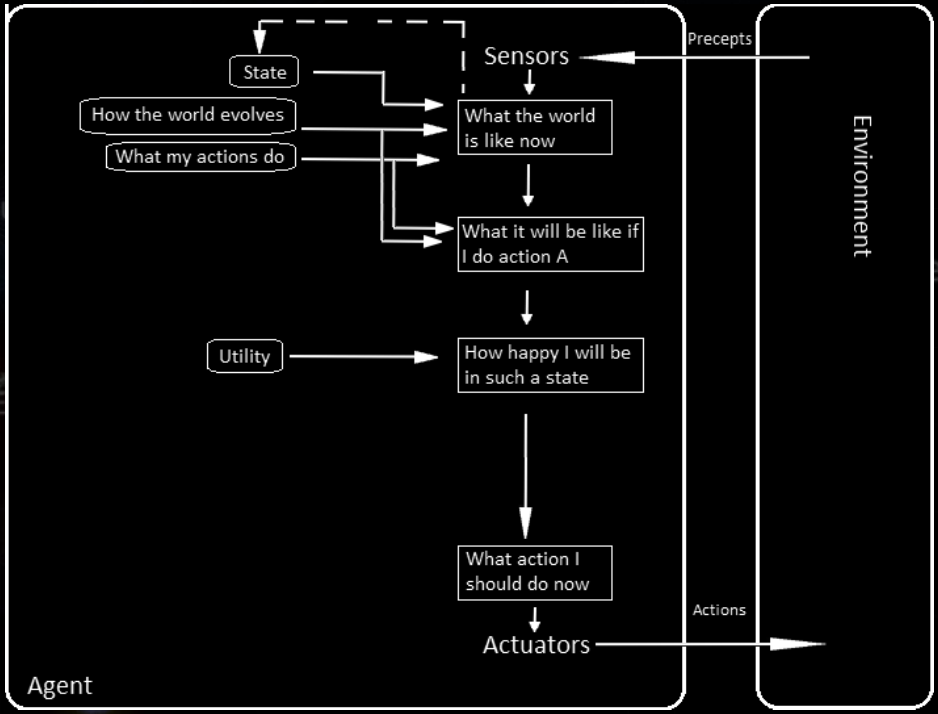
- Why do we need them?
  - Communication latency with interstellar spacecraft
  - Intelligence requiring tasks
    - Navigation
    - Scientific experiments
    - Emergencies / Fault-tolerance
- Advantages:
  - Eliminate significant mass for
    - Habitat module
    - Extra shielding
  - Longevity
  - Rationality
  - Fast reaction time
  - Error recovery

100 YEAR STARSHIP™
sigma 2398 11.5 LY

Agent Designs



State
How the world evolves
What my actions do
Sensors
What the world is like now
What it will be like if I do action A
Goals
What action I should do now
Actuators
Agent
Precepts
Actions
Environment



State
How the world evolves
What my actions do
Sensors
What the world is like now
What it will be like if I do action A
Utility
How happy I will be in such a state
What action I should do now
Actuators
Agent
Precepts
Actions
Environment

100 YEAR STARSHIP™
sigma 2398 11.5 LY

Artificial General Intelligence

- AGI: human-level general-purpose AI
  - Solomonoff's universal induction theory
    - Assumption: probability distribution of the universe is computable
      - Holds for the observable universe due to Bekenstein bound
    - Converges very rapidly (convergence theorem)
      - Predictions with small error dependent on universal computer
    - Incomputable, but can be approximated
  - Probability of bitstring x
    - $$P(x) = \sum_{U(p_i)=x*} 2^{-|p_i|}$$
    - Applied to prediction: sequence, set, operator prediction
    - Fully formalizes Occam's razor and Epicurus's law of multiple explanations
    - Universal prior for any application of Bayes theorem

AGI: How Does That Work?

- Operator induction:
  - Given set $\{(x_1, y_1), (x_2, y_2), \dots\}$ of input,output pairs
  - Learn generalized conditional pdf P(Y|X)
  - We can now predict $argmax_Y P(Y|X)$
  - Operator induction solves any classical ML problem
  - You can apply induction to solve any AI problem
- Approximation methods
  - Universal Search (Levin, 1973)
  - Evolutionary Programming (Nature, since big bang)
- Application
  - Cognitive architecture that uses inductive inference module

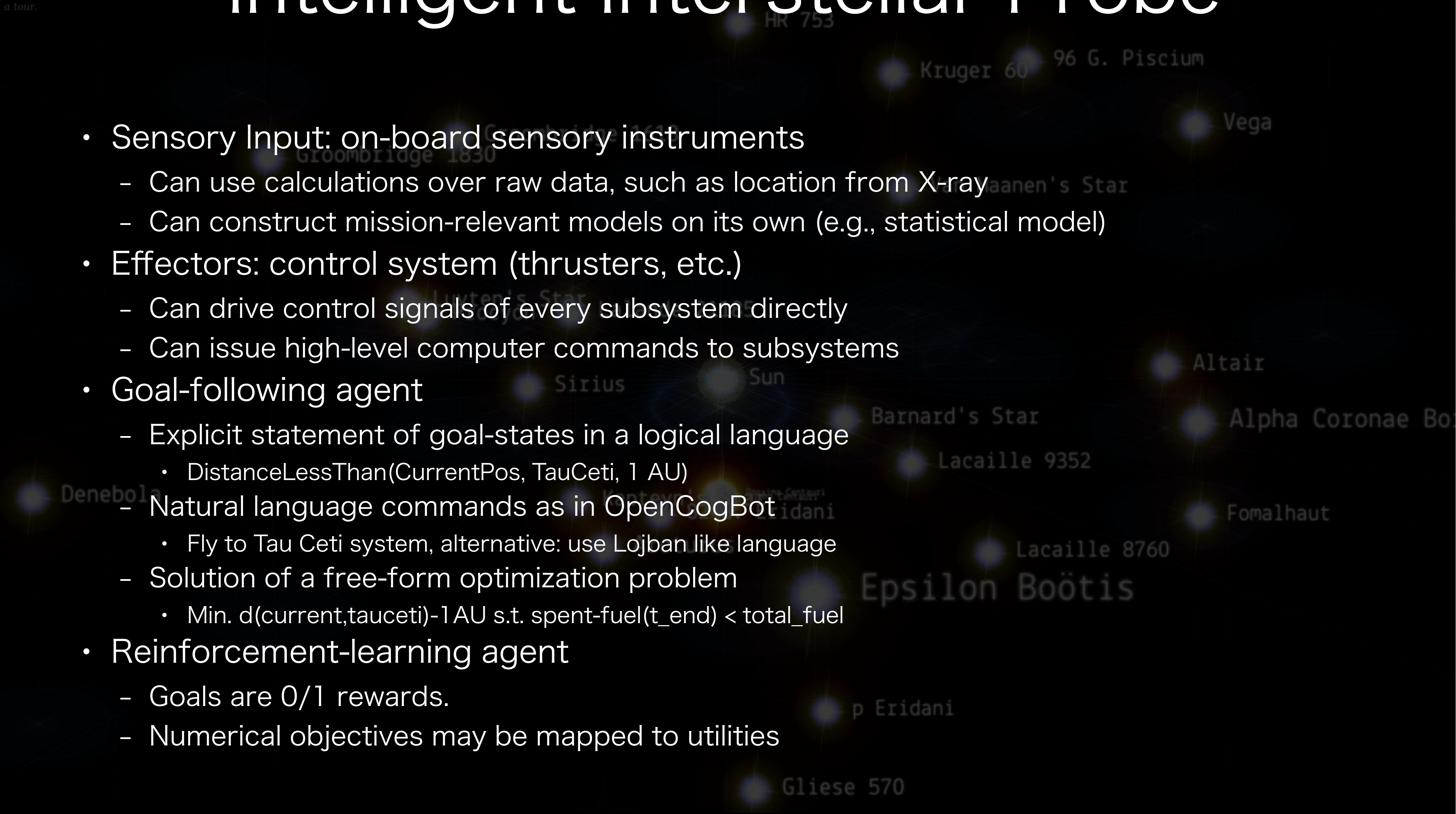
Some Challenges of AGI Research

- Choosing the right universal computer
  - Variants of LISP, machine language, MATLAB, etc.
- Efficient approximation algorithm
  - Search space is exponential in the number of bits
  - Parallelism
- Memory:
  - Design long-term memory (transfer learn.)
  - How to make memory work with induction?
- Modularity and Scalability

AGI Systems

- Alpha (Solomonoff, 2002):
  - Assimilates narrow-AI systems, modular
  - Higher-order cognitive procedures (analysis, synthesis)
  - Can solve free-form time-limited optimization problems
- Gödel Machine (Schmidhuber, 2003):
  - Self-reflective AI agent design
  - Can use any reaction policy
- AIXI (Hutter, 2006):
  - Extends universal sequence induction
  - Optimal reinforcement learning agent model
    - input, actions, rewards
    - Can optimize expected future rewards

Intelligent Interstellar Probe



- Sensory Input: on-board sensory instruments
  - Can use calculations over raw data, such as location from X-ray
  - Can construct mission-relevant models on its own (e.g., statistical model)
- Effectors: control system (thrusters, etc.)
  - Can drive control signals of every subsystem directly
  - Can issue high-level computer commands to subsystems
- Goal-following agent
  - Explicit statement of goal-states in a logical language
    - DistanceLessThan(CurrentPos, TauCeti, 1 AU)
  - Natural language commands as in OpenCogBot
    - Fly to Tau Ceti system, alternative: use Lojban like language
  - Solution of a free-form optimization problem
    - Min. d(current,tauceti)-1AU s.t. spent-fuel(t_end) < total_fuel
- Reinforcement-learning agent
  - Goals are 0/1 rewards.
  - Numerical objectives may be mapped to utilities

100 YEAR STARSHIP™
sigma 2398 11.5 LY

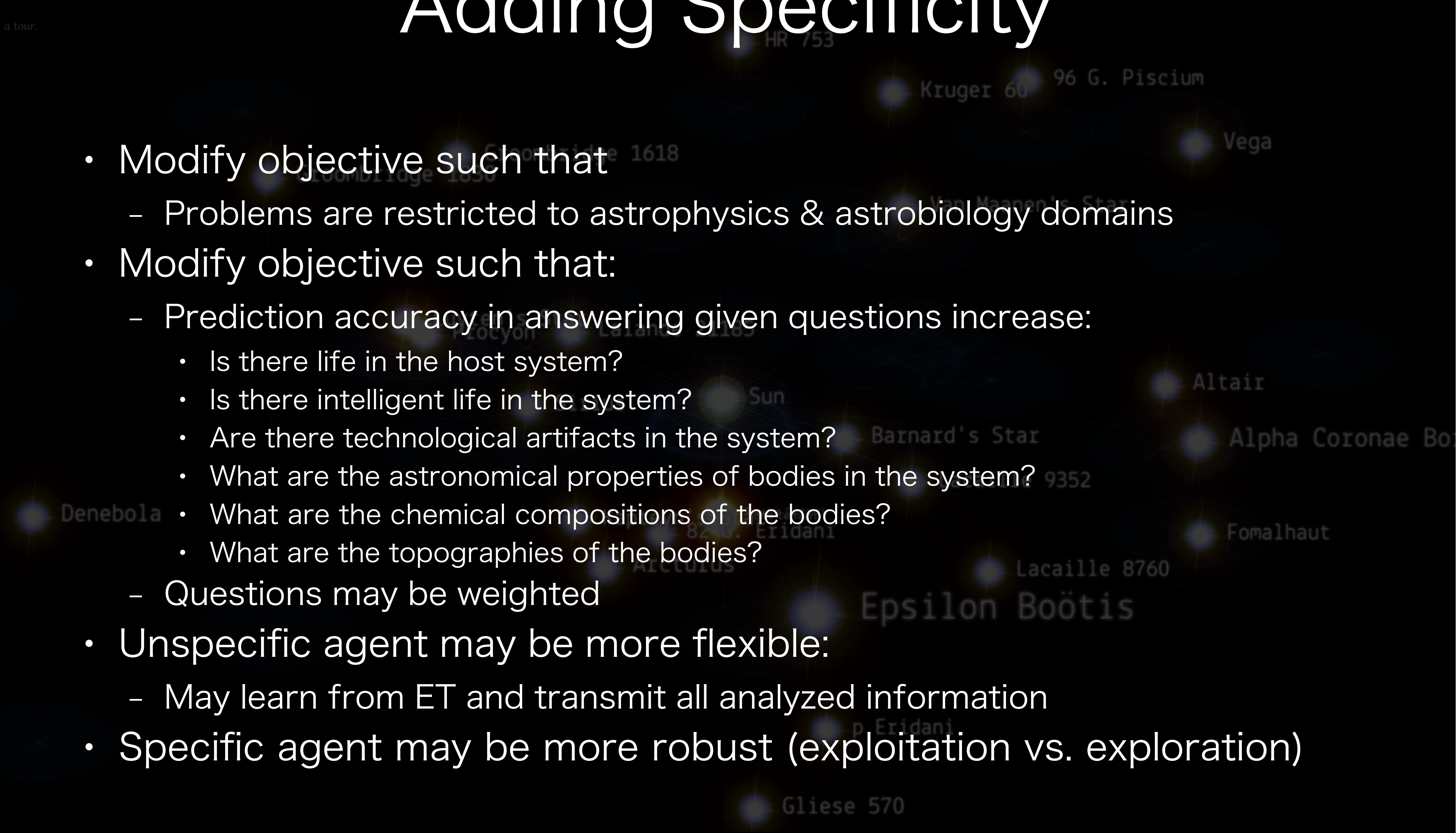
Autonomy

- Specific vs. unspecific goals:
  - Specific goal: numerically precise, model states
    - Example: Orbit around target star at 1.0 AU +/- 10 m dist
    - Runs risk of failure, due to tight constraints
  - Unspecific goal: allow freedom to plan
    - Example: gather as much information as possible!
    - We call this a universal goal
    - Universal goals are likely more appropriate for intelligent agents
- Fully-autonomous vs. semi-autonomous agent:
  - Fully autonomous agents diverge and have drives
  - We propose simple solutions for this:
    - Use physical constraints
    - Specify the goals generally, but precisely
    - Avoid open-ended goals

Autonomous Agent Example

- Can be expressed as an optimization problem
  - Max. amount of (true) information discovered & transmitted back to Earth (cumulative, for all future actions)
    - With respect to all questions that may be asked of the system
      - May formalize as improving prediction accuracy for any induction problem
    - Weighted according to a priori probability of question
    - RL formulation: knowledge-seeking agent (loosely equivalent)
  - Such that
    - Spacecraft resources are not depleted
    - Spacecraft is operational for specified period
    - Probe remains within given space-time region, energy limits
      - Primary Directive? Generalized non-interference clause
- Advanced formulation (Alpha Stage 2)
  - Optimizes efficiency of knowledge extraction bits/J.sec

Adding Specificity



- Modify objective such that
  - Problems are restricted to astrophysics & astrobiology domains
- Modify objective such that:
  - Prediction accuracy in answering given questions increase:
    - Is there life in the host system?
    - Is there intelligent life in the system?
    - Are there technological artifacts in the system?
    - What are the astronomical properties of bodies in the system?
    - What are the chemical compositions of the bodies?
    - What are the topographies of the bodies?
  - Questions may be weighted
- Unspecific agent may be more flexible:
  - May learn from ET and transmit all analyzed information
- Specific agent may be more robust (exploitation vs. exploration)

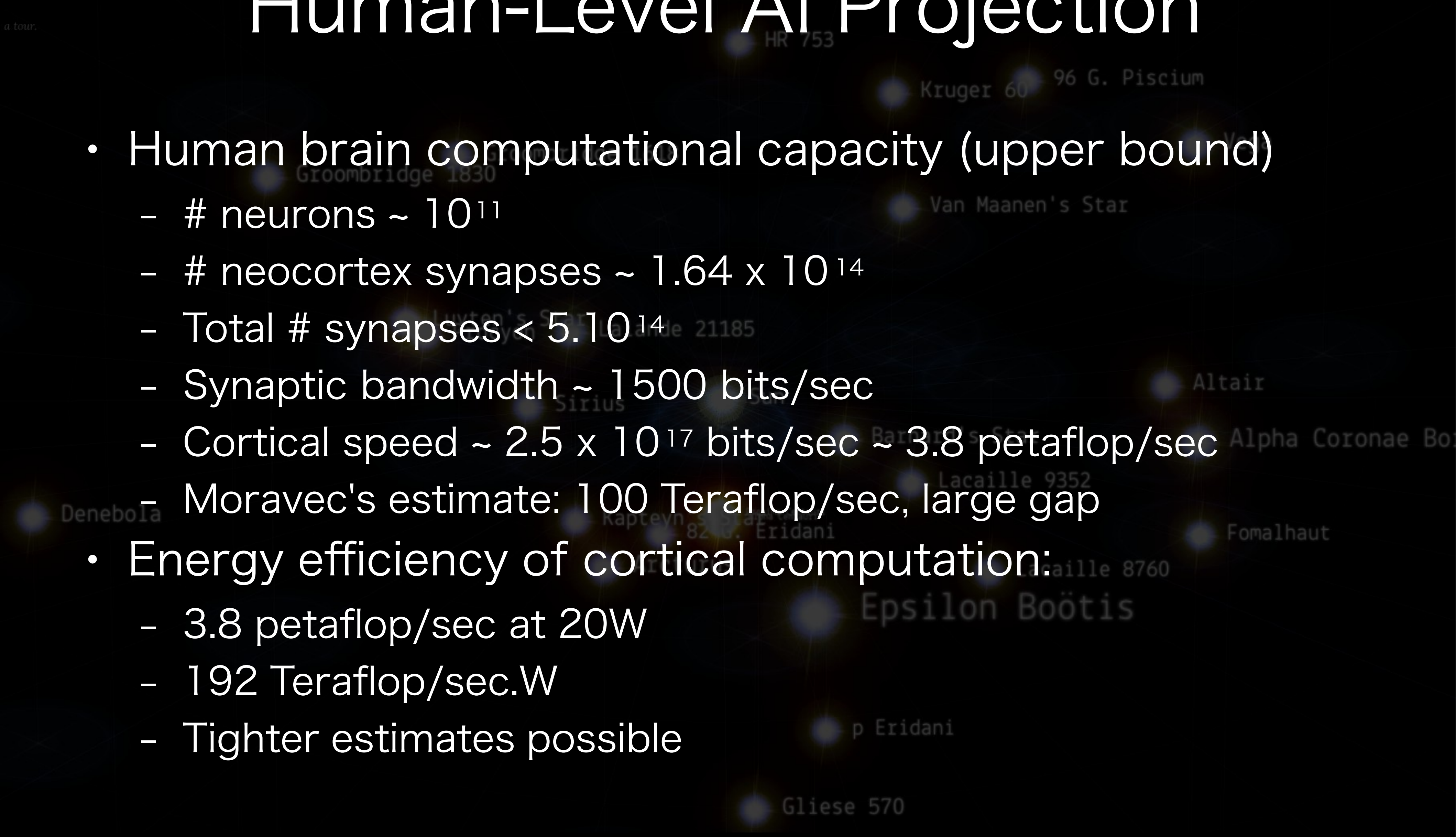
Intelligent Probe Training

- Various modes of intelligence may be sought
  - A high-level gator system that acts like an animal (instinctive)
  - A human-level AI that has expertise in related fields
  - A self-reflective, self-improving trans-sapient agent:
    - “Artificial starfleet captain”
- Starts in 2030 using supercomputers
- Most expensive part of an AGI application

Intelligent Probe Training

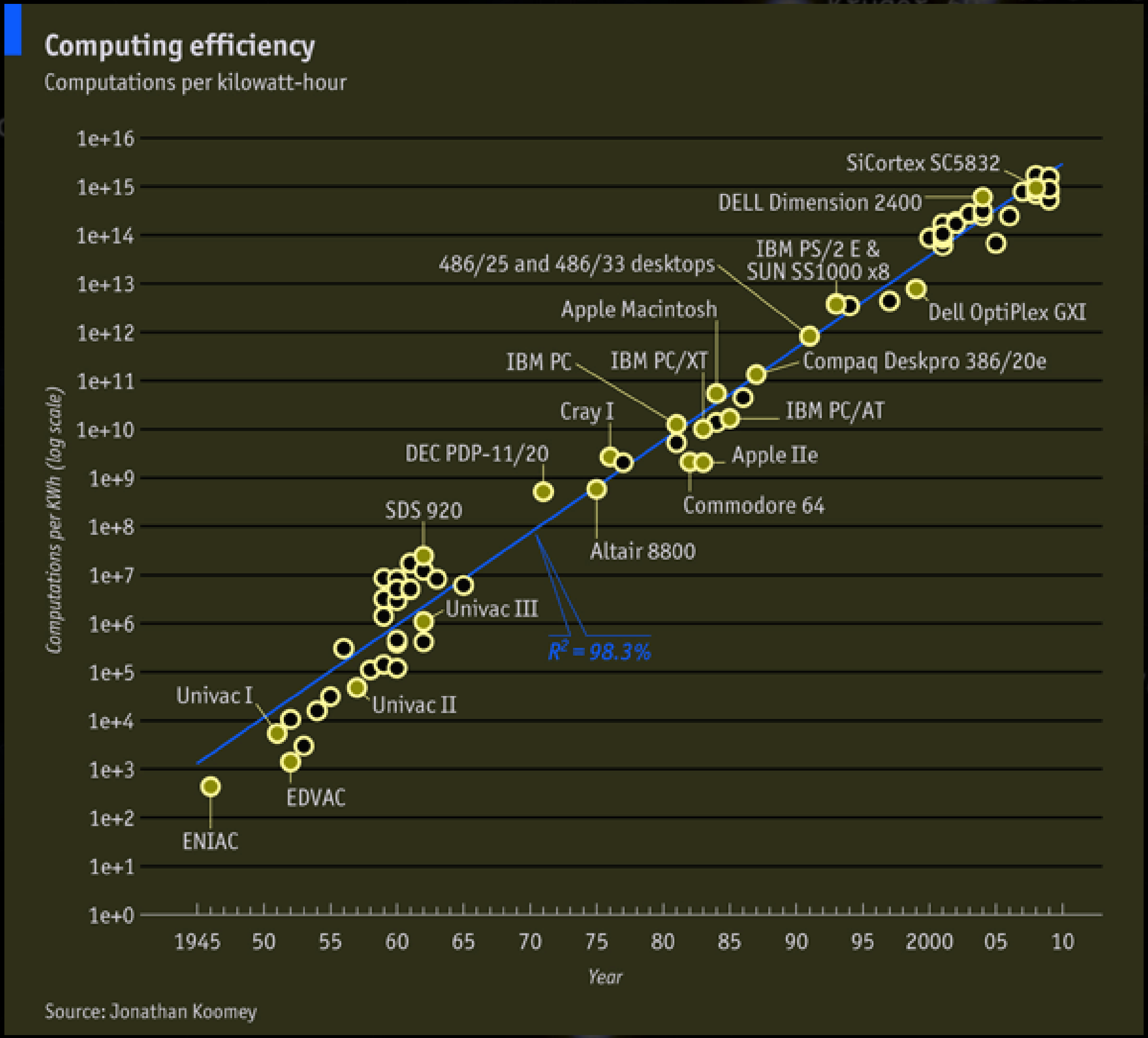
- Background knowledge & testing
  - Standard material in any human-level AI
    - Training in pattern recognition problems (visual, audio, etc.)
    - Extensive training in math, physics, English
    - Includes much of high-school/college curriculum
    - Capability to learn and apply human knowledge assessed (can pass SAT/GRE)
  - Interstellar mission knowledge
    - Expert-level training in physics, chemistry, astrophysics, astronomy, astrobiology
    - Spacecraft specific training for use and control of every subsystem
    - Mission training involving hand-crafted and random mission simulations
    - Mission readiness determined by an Interstellar Turing Test

Human-Level AI Projection



- Human brain computational capacity (upper bound)
  - # neurons ~ 10^{11}
  - # neocortex synapses ~ 1.64×10^{14}
  - Total # synapses < 5.10^{14}
  - Synaptic bandwidth ~ 1500 bits/sec
  - Cortical speed ~ 2.5×10^{17} bits/sec ~ 3.8 petaflop/sec
  - Moravec's estimate: 100 Teraflop/sec, large gap
- Energy efficiency of cortical computation:
  - 3.8 petaflop/sec at 20W
  - 192 Teraflop/sec.W
  - Tighter estimates possible

Koomey's Law



Computing efficiency

Computations per kilowatt-hour

Source: Jonathan Koomey

Human-Level AI Projection

- Koomey's law:
  - energy-efficiency doubles every 18 months:
  - 2026: Human-level energy efficiency (Ozkural 2011)
- Current specs (2013):
  - $100 → 32 GFlops (Adapteva)
  - 72 GFlops/watt (Adapteva 64-core CPU)
- Extrapolating Koomey's law for upper bound:
  - 2013 + 17 years ~ 2030
- By then human-level AI will already be available

Infinity Point Hypothesis

- Solomonoff's analysis of social effects of AI (1985)
- Human-level AI accelerates Moore's law (1985)
  - Moore's law:
    - Number of transistors placed on a microprocessor at a fixed cost doubles every two years
    - Current doubling time is about three years (!)
  - CS community size ~ rate of improvement
  - CS community size ~ rate of log of computing efficiency
    - The larger the community, the smaller the doubling time gets
  - Fixed amount of money invested in AI every year
  - Infinite improvement in finite time! (Math. singularity)
    - In practice we have only finite improvement, 600 yrs → few decades
  - We can replace Moore's law with Koomey's law

Brain Simulation for Virtual Crew

- Brain scanning tech improves fast
  - Human Brain Project will complete by 2025
- By 2040, 20W can simulate ~100 virtual crew
  - We can build a “spaceship of the imagination”
- As Tipler and others suggested
  - Sims interface with body images in VR
  - Sims can occupy a “virtual bridge” and a virtual city
  - Sims can carry a copy of the internet with them
- Trans-sapient artificial astro-physicist:
  - Extended cortex in sim for astrophysics modalities
  - Can invoke computer tools with no delay
- Sims will be trained before voyage on virtual missions

Computer Technology by 2040

- Without infinity point:
  - 195 Petaflop/sec.W efficiency (101.5 x human)
  - 1 brain simulation ~ 0.2W
- With infinity point:
  - Start from 2030, assume human-level AI
  - At 20W per sim, we double CS community every year
    - Negligible cost w.r.t. global economy
  - Super-exponential improvement in energy efficiency
  - Infinity point then is reached in 4.62 years
    - Computational resources are abundant by 2040

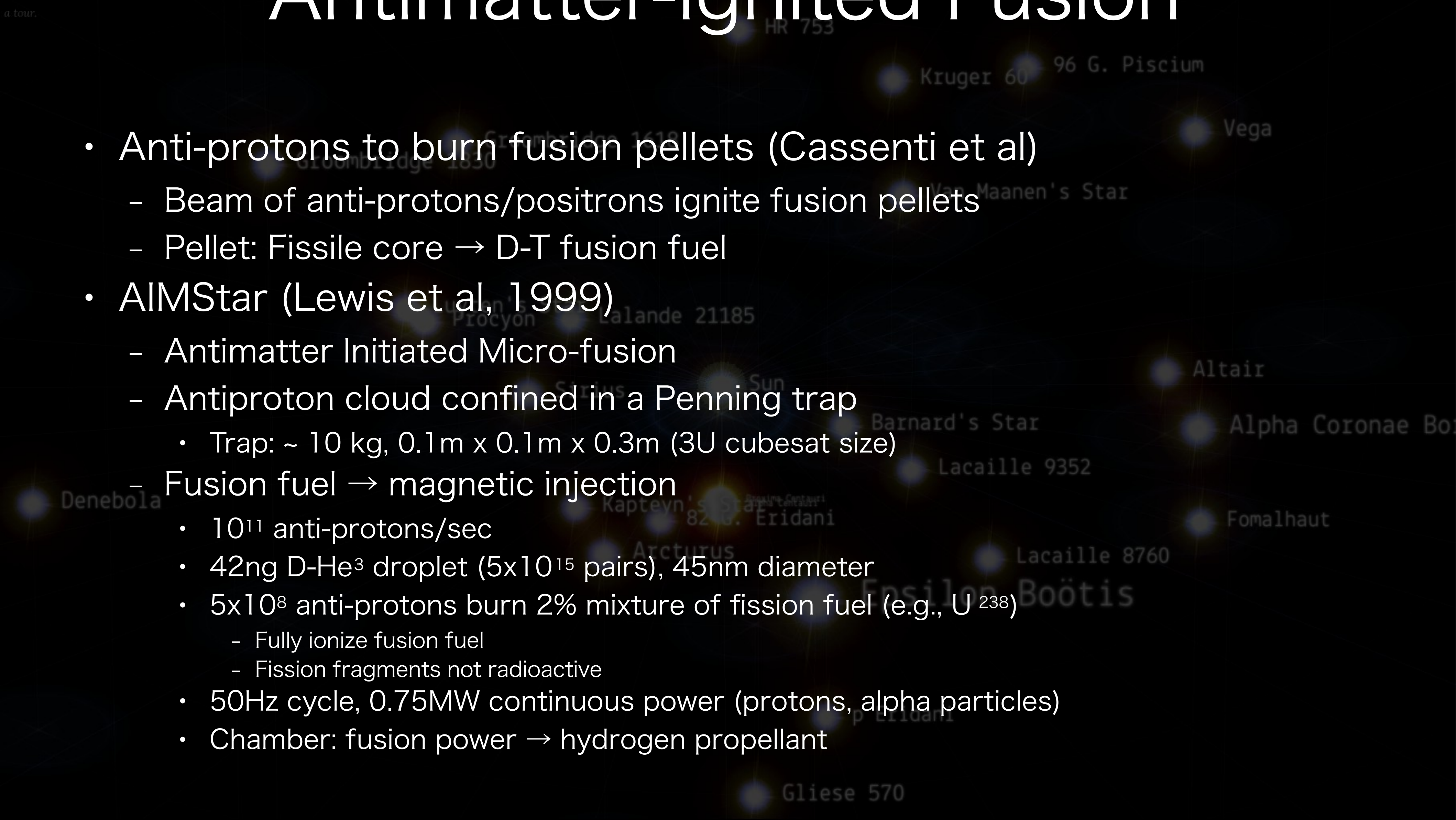
Command and Control Module

- Ample speed to simulate 100 humans
- Allow for x5 resources for VR and programs
- Or 1 trans-sapient AGI agent (x500 human)
- 19.5 exaflop/sec speed
- 10 exabyte storage
- Three-way redundancy
- 100/300 W power, 1kg payload (1U unit)
- Can scale down to 2/6W on-route (10x human)
  - Long voyage, after all

Propulsion

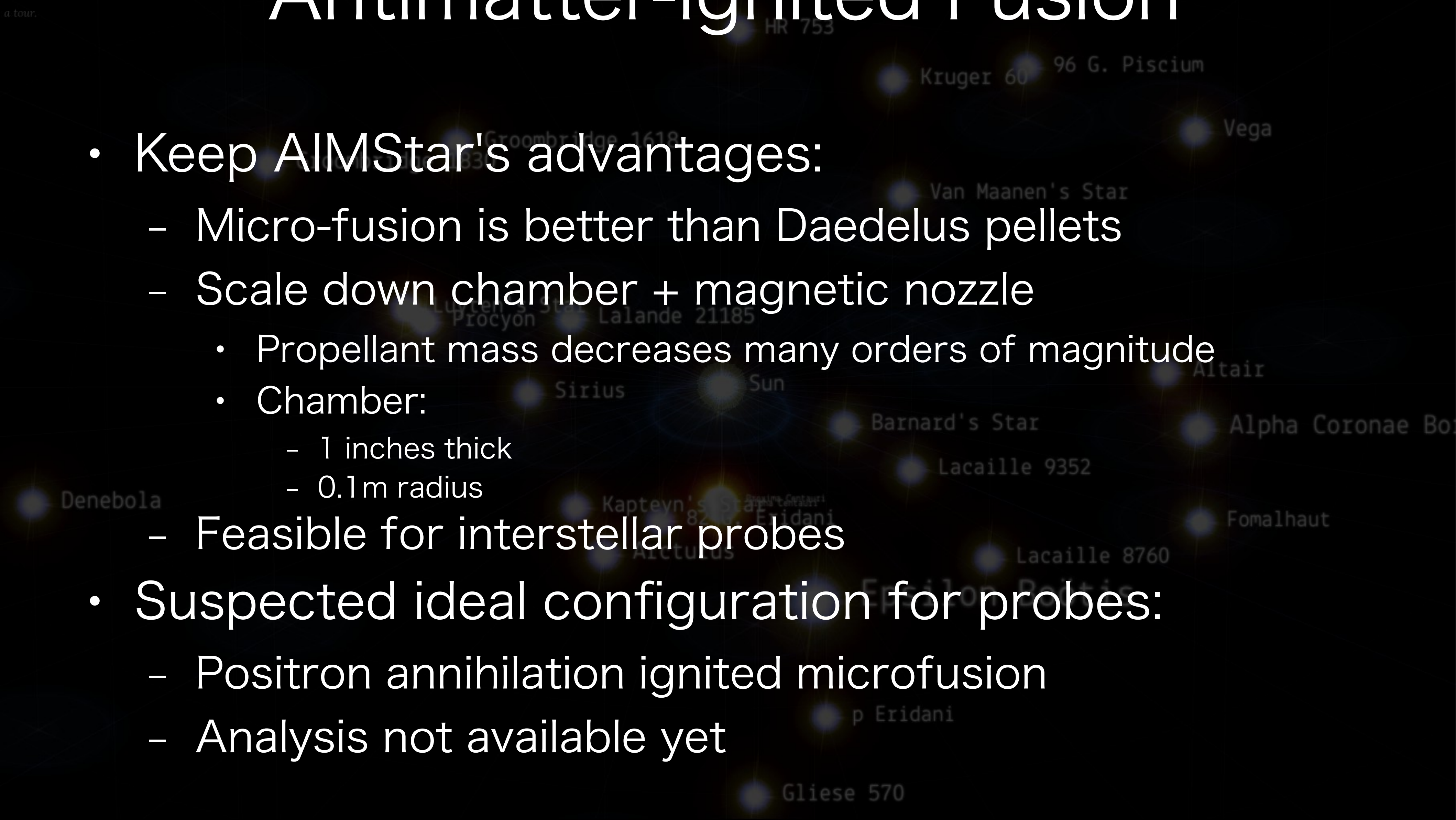
- Fission, fusion, anti-matter and sail concepts feasible
- Fission/fusion too massive for small probes
  - Laser/beam driver is massive, requires auxiliary power
  - Minimum pellet size limitation, massive containment
- Light-sail inappropriate for rendezvous missions
- Pure antimatter rocket requires much antimatter
- Anti-proton / Positron ignited fusion
  - Use antimatter annihilation to ignite D-He3 fusion
  - Overcomes driver overhead
  - We re-evaluate AIMStar and Positron rocket studies
    - AIMStar uses antiproton annihilation for a cubesat mission
    - Positronics Research LLC is developing positron rocket
    - Positron Dynamics LLC is working on positron storage

Antimatter-Ignited Fusion



- Anti-protons to burn fusion pellets (Cassenti et al)
  - Beam of anti-protons/positrons ignite fusion pellets
  - Pellet: Fissile core → D-T fusion fuel
- AIMStar (Lewis et al, 1999)
  - Antimatter Initiated Micro-fusion
  - Antiproton cloud confined in a Penning trap
    - Trap: ~ 10 kg, 0.1m x 0.1m x 0.3m (3U cubesat size)
  - Fusion fuel → magnetic injection
    - 10^{11} anti-protons/sec
    - 42ng D-He^3 droplet ($5x10^{15}$ pairs), 45nm diameter
    - $5x10^8$ anti-protons burn 2% mixture of fission fuel (e.g., U^{238})
      - Fully ionize fusion fuel
      - Fission fragments not radioactive
    - 50Hz cycle, 0.75MW continuous power (protons, alpha particles)
    - Chamber: fusion power → hydrogen propellant

Antimatter-Ignited Fusion



- Keep AIMStar's advantages:
  - Micro-fusion is better than Daedelus pellets
  - Scale down chamber + magnetic nozzle
    - Propellant mass decreases many orders of magnitude
    - Chamber:
      - 1 inches thick
      - 0.1m radius
  - Feasible for interstellar probes
- Suspected ideal configuration for probes:
  - Positron annihilation ignited microfusion
  - Analysis not available yet

Interstellar Communication

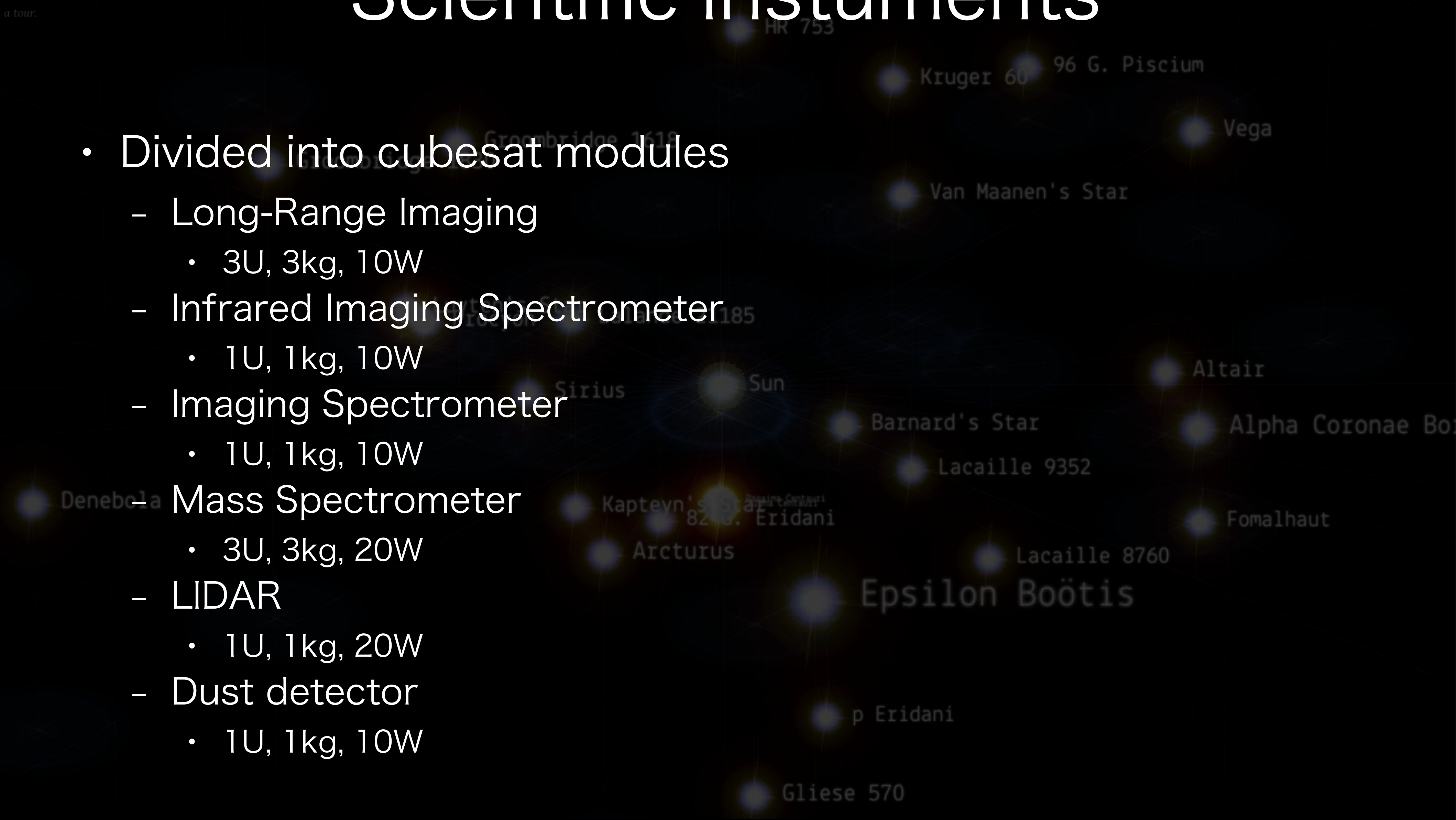
- Two reasonable approaches
  - Laser communication
    - 20W and a 3 meter telescope like Hubble
  - Gravitational-lensing amplification for RF (FOCAL)
    - Exploits gravitational-lensing of Sun (and host system)
    - Requires a mission to 550AU distance from Sun
  - Gravitational-lensing preferable
    - 40 W power
    - Uses only an inflatable 18m RF antenna → small mass
  - Inflatable 20m RF Antenna?
    - 3U, 5kg, 30W
    - 8m parabolic dish
    - Mylar / Meta-materials → unsolved yet (!)

100 YEAR STARSHIP™
sigma 2398 11.5 LY

Navigation

- Four pulsars can be used to compute location (Deng et al, 2013)
- X-Ray telescope is required
  - Traditional x-ray telescopes too heavy
  - Lobster eye optics improves efficiency up to 1000x
  - Mini lobster X-ray imaging module for picosatellites (Hudec et al):
    - 10cm x 10cm x 30cm design underway (3U cubesat)
    - We estimate 3kg device with 10W power
- We also need
  - Attitude control, auxiliary thrust
    - 3 x 3U CAT plasma thruster modules, 20W, 5 kg
    - 2mN continuous thrust for 10W, 20mN thrust for pulsed 100W
    - 2 modules have 4x2 small thrusters, 1 module 3-axis 5x2
  - Sensor module
    - 3-axis: accelerometers, gyros, magnetometers
    - Attitude-meter, star-sensor
    - Radiation monitor
    - 1U, 1kg, 10W

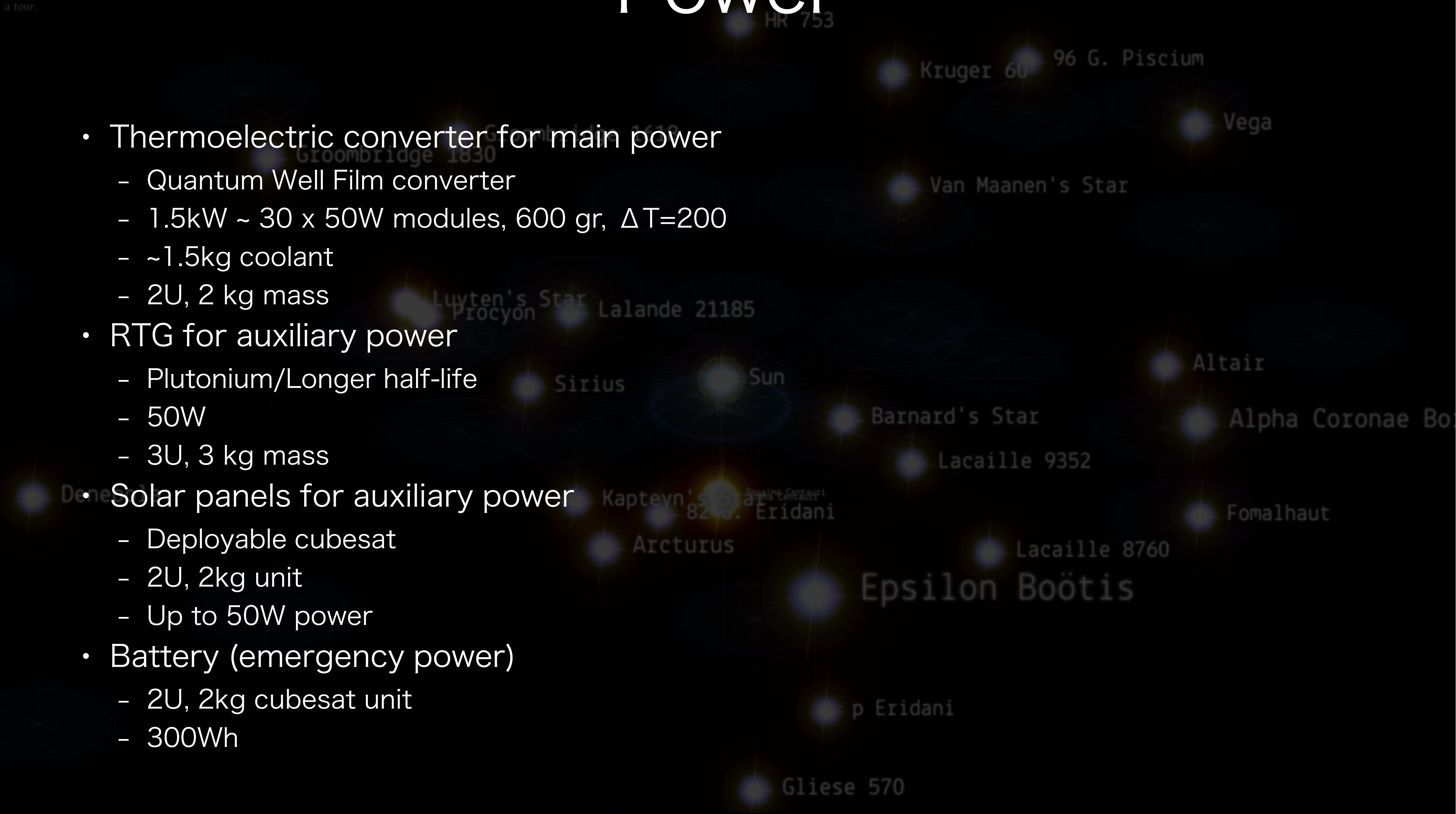
Scientific Instuments



- Divided into cubesat modules
  - Long-Range Imaging
    - 3U, 3kg, 10W
  - Infrared Imaging Spectrometer
    - 1U, 1kg, 10W
  - Imaging Spectrometer
    - 1U, 1kg, 10W
  - Mass Spectrometer
    - 3U, 3kg, 20W
  - LIDAR
    - 1U, 1kg, 20W
  - Dust detector
    - 1U, 1kg, 10W

100 YEAR STARSHIP™
sigma 2398 11.5 LY

Power



- Thermoelectric converter for main power
  - Quantum Well Film converter
  - 1.5kW ~ 30 x 50W modules, 600 gr, $\Delta T=200$
  - ~1.5kg coolant
  - 2U, 2 kg mass
- RTG for auxiliary power
  - Plutonium/Longer half-life
  - 50W
  - 3U, 3 kg mass
- Solar panels for auxiliary power
  - Deployable cubesat
  - 2U, 2kg unit
  - Up to 50W power
- Battery (emergency power)
  - 2U, 2kg cubesat unit
  - 300Wh

100 YEAR STARSHIP™
sigma 2398 11.5 LY

Shielding

- Artificial mini-magnetosphere (Bamford, 2008)
  - Imitates Earth's magnetosphere
  - Generated magnetic field
    - Lab experiment used 0.5T natural magnet
    - 50nT sufficient for space applications
    - ~100m radius sufficient
  - New versions require little power
    - We estimate 500W, produced from fusion burn
    - We estimate 1kg additional weight for mini-shield
- Carbon nanotube EMI shield
  - Previously tested on cubesats

Subsystems Summary

Name	Mass (kg)	Power (W)
Propulsion		
Reaction Trap	10	100
Chamber	32	
Fusion fuel		
Power		
Thermoelectric	2	
RTG	3	
Solar Panel	2	
Battery	2	
Navigation		
Mini-Lobster	3	10
Sensors	1	10
Attitude Ctrl.	15	60

Name	Mass (kg)	Power (W)
Science		
Long-Range	3	10
IR Imaging	2	10
Imaging Spectr.	2	10
Mass Spectr.	3	20
LIDAR	1	20
Dust Detector	1	10
Control		
AI	1	2-300
Shielding	1	0-500
Communication		
Inflatable RF	5	30
Spaceship Infr.	2	10
Total	91	302-1102

100 YEAR STARSHIP™
sigma 2398 11.5 LY

Mission Profile

	Tau Ceti	*Gliese 667C*
Distance	11.9 lightyears	22 lightyears
Cruise Speed	0.1c	0.1c
Burnout time	9.183 years	9.183 years
Burnout distance	0.258 lightyears	0.258 lightyears
Antimatter required	0.97 milligrams	0.97 milligrams
Fusion fuel required	36469 kg	36469 kg
Cruise Time	113.8 years	214.8 years
Total Voyage Time	132.2 years	233.2 years

Mission Profile for Q-Thruster

	Tau Ceti	*Gliese 667C*
Distance	11.9 light years	22 light years
Mass	1561 kg	1561 kg
Cruise Speed	0.5c	0.5c
Burnout time	18.5 years	18.5years
Burnout distance	4.63 light years	4.63 light years
Cruise Time	5.26 years	25.4 years
Total Voyage Time	42.3 years	62.5 years

- SAFE-400 Reactor
  - 512kg
  - 400kW / 100kWe
- We assume 4N/kW performance, 1000kg q-thruster

References

- RR de Ruyter van Steveninck & SB Laughlin,*The rate of information transfer at graded potential synapses*, Nature 379, 642-645 (1996).
- Eray Özkural. *Diverse Applications of Algorithmic Probability Theory.* Solomonoff Memorial Conference, Melbourne, Australia (2011)
- Ray Solomonoff, 1985 *The Time Scale of Artificial Intelligence: Reflections on Social Effects*, Human Systems Management, Vol. 5, pp. 149-153, 1985
- Koomey, J.G., Berard, S., Sanchez, M., Wong, H.: *Implications of historical trends in the electrical efficiency of computing*. IEEE Annals of the History of Computing 33 (2011) 46–54
- Solomonoff, R.J.: *A formal theory of inductive inference, part i.* Information and Control 7(1) (1964) 1–22

References (cont.)

- Levin, L.A.: *Universal sequential search problems*. Problems of Information Transmission 9(3) (1973) 265–266
- Solomonoff, R.J.: *Algorithmic Probability – Its Discovery – Its Properties and Application to Strong AI.* In: Randomness Through Computation: Some Answers, More Questions. World Scientific Publishing Company (2011) 149–157
- Schmidhuber, J.: *Ultimate cognition a la Godel.* Cognitive Computation 1(2) (2009) 177–193
- Hutter, M.: *Universal algorithmic intelligence: A mathematical top→down approach.* In Goertzel, B., Pennachin, C., eds.: Artificial General Intelligence. Cog- nitive Technologies. Springer, Berlin (2007) 227–290